1982

The Full Optimized Reaction Space model for quantum chemical reaction calculations. Definition, applications, and IntraAtomic Correlation Correction extension

Michael William Schmidt
Iowa State University

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Schmidt, Michael William

THE FULL OPTIMIZED REACTION SPACE MODEL FOR QUANTUM CHEMICAL REACTION CALCULATIONS. DEFINITION, APPLICATIONS, AND THE INTRAATOMIC CORRELATION CORRECTION EXTENSION

Iowa State University PH.D. 1982

The Full Optimized Reaction Space model for quantum chemical reaction calculations. Definition, applications, and the IntraAtomic Correlation Correction extension

by

Michael William Schmidt

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Chemistry
Major: Physical Chemistry

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

I. INTRODUCTION

A strong impact of quantum chemistry can be expected to occur in the field of chemical reactions because theoretical calculations of reaction paths and transition states are, in principle, no different in nature from those of stable molecules, whereas the experimental elucidation of reaction intermediates is subject to many uncertainties due to their fleeting appearance and disappearance.

However, even for the theoretical approach, there exists a marked difference between the calculation of stable species and that of molecular structures characteristic for most reaction intermediates. While the standard self-consistent-field method has proved to be a very serviceable "dominant approximation" for many stable molecules, it is rarely adequate for systems in the flux of reactive changes, where orbitals not only deform but, in addition, change occupation numbers due to changes in configurational mixing. In fact, on the basis of currently available experience, it is not possible to anticipate which configurations will dominate a wavefunction at various points on a reactive path. Consider, for example, the case that reactants and products are well described by single determinant SCF functions and that many, but not all of the occupied orbitals continuously deform into certain occupied product orbitals. Some of the doubly occupied reactant orbitals deform, however, into virtual (unoccupied) product orbitals and, correspondingly, some of the virtual reactant orbitals deform into doubly occupied product orbitals. In such a case, it is tempting to calculate the reaction using a two determinant wavefunction. But more often than not

in the intermediate region where the "reaction orbitals" have occupation numbers close to unity, there exist numerous other configurations in addition to the reactant and product configurations which have similar importance in the wavefunction. Their inclusion in the dominant part of the wavefunction proves to be essential, in particular, for the calculation of reaction barriers.

Conversely, it is also important to know which parts of the exact wavefunction can be completely neglected for calculations along a specific reaction path. Since full recovery of the correlation energy by direct computation is out of the question, even with present day computers, the practical goal of quantum chemistry is always the calculation of energy curves which are reliably parallel to the exact energy surfaces by including all those correlation effects which change significantly (i.e., more than kT) while omitting those effects which do not and, hence, would cancel when energy differences are taken. This parallelism is required for quantitative evaluation of both overall reaction energetics and barrier heights.

From the preceding remarks, it can be inferred that it is essential to avoid unjustified restrictions in the selection of configurations or in the orbitals from which these configurations are formed, so that all pertinent changes can be reflected without bias in the calculation. As mentioned, this is even more crucial for transition states than products or reactants. Unbiased configuration generating orbitals can be guaranteed by use of a multi-configurational SCF approach with adequately large atomic basis sets. Elimination of prejudice in the

selection of configurations calls for an MCSCF model that makes provision for the free interaction of all configurations which may become significant during the course of the molecular rearrangements. It is, therefore, necessary to formulate a method for determining configuration generating orbitals and for selecting configurations which is general, unambiguously defined, without any bias, and quantitatively reliable. In addition, it is desirable that the model lend itself in a natural, but rigorous way, to suitable chemical interpretation in terms of atoms and bonds.

The Full Optimized Reaction Space (FORS) model for the treatment of chemical reactions is described in Chapter II. The FORS model is illustrated, and its quantitative reliability assessed by calculations on diatomic molecules. In Chapter III, certain orbital invariances in FORS wavefunctions are exploited to obtain a number of different, yet equivalent sets of orbitals, including a new type of localized orbital, obtained by projection. These orbital sets reveal different chemical aspects of the FORS wavefunction, including molecular bonding and the participation of atoms in the molecular environment. Two polyatomic reactions are examined in the next two chapters. In Chapter IV, the dihydrogen exchange reaction between ethane and ethylene is shown via SCF calculations to possess a large reaction barrier, in spite of the reaction's symmetry-allowedness. FORS calculations serve to reinforce this result. Various orbital sets from both the SCF and FORS wavefunctions are used to analyze the presence and magnitude of the barrier to this exchange. Chapter V deals with a reaction important in gas phase

ozonolysis. The insertion of singlet methylene into singlet oxygen to form the small ring molecule dioxirane, the opening of this ring to dioxymethane, and the decomposition of dioxymethane to hydrogen and carbon dioxide are examined. This reactive sequence features extensive configurational mixing and, hence, is treated entirely at the FORS level. Finally, Chapter VI describes how a FORS wavefunction can be transformed, with the help of the projected localized orbitals, into an equivalent function that is a superposition of products of atomic and ionic states. Such an analysis reveals how various atomic and ionic valence states participate in molecular wavefunctions. Furthermore, knowledge of such "atoms-in-molecules" expansions permits the incorporation of data from atomic spectra according to an empirical correction scheme termed the IntraAtomic Correlation Correction. This correction to the FORS wavefunction is shown to yield improved results for the bond strengths of diatomics.

A number of abbreviations will be used repeatedly in the following chapters. Many of these abbreviations are common quantum chemical usage, while others may not be. These terms are defined where first used, but are gathered here for convenience.

AO Atomic Orbital - The functions, usually atom centered, in which the molecular orbitals are expressed.

PAO Primitive Atomic Orbital - An exponential or Gaussian AO with fixed exponent, multiplied by appropriate angular dependence to form an s, p, d, etc. PAO.

QBO Quantitative Basis Orbital - A fixed linear combination of PAOs. There are ordinarily fewer QBOs than PAOs. These "contractions" combine only functions with the same angular dependence on the same nucleus. The MOs are expanded in the QBO basis.

MBS Minimal Basis Set - If the number of QBOs used on an atom just equals the number of orbitals occupied in that atom's SCF function, the basis is minimal. Ordinarily an extended, more flexible, basis is used.

MO Molecular Orbital - A linear combination of QBOs. There are a number of types:

CGO Configurational Generating Orbital - Any MO which is occupied in any of the configurations in the wavefunction.

NRO Natural Reaction Orbitals - The particular set of reactive MOs whose first order density matrix is diagonal. Often called simply NOs.

LRO Localized Reaction Orbitals - Any set of reactive MOs which has been localized by any feasible method.

PLO Projected Localized Orbitals - A set of core and reactive MOs which has been localized by a projective technique. The PLOs are usually very atomic in character.

The above types of AOs and MOs depend on the coordinates of only a single electron. The following terms refer to N electron basis functions:

SAAP Spin Adapted Antisymmetrized Product - A product of space orbitals (MOs), multiplied by a Serber spin function, and antisymmetrized.

CSF Configurational State Function - Often identical with a SAAP. In cases of high point symmetry, a CSF is a fixed linear combination of a few SAAPs having the correct spatial symmetry.

CF Composite Function - An antisymmetrized product of atomic eigenstates. This is a different type of N electron function than a CSF.

There are three major types of MO calculations:

SCF Self-Consistent Field - A type of wavefunction using only one configuration (CSF) for which the expansion of the MOs in the QBO basis is optimized.

CI Configuration Interaction - A type of wavefunction in which more than one configuration is used, but the MOs from which these configurations are built are held constant.

MCSCF Multi-Configuration SCF - The "logical or" of SCF and CI. Both the expansion coefficients of the configurations, and the expansions of the MOs in the QBO basis are optimized.

A few other terms are given below.

FRS Full Reaction Space - A set of all possible CSFs (or CFs) obtained by occupying the reactive orbitals by the reactive electrons in all possible ways.

FORS Full Optimized Reaction Space - A MCSCF model, wherein the orbitals used to generate a FRS are optimized for that particular configurational space via MCSCF calculation.

AIM Atoms in Molecules - An empirical correction scheme, whereby a molecular wavefunction is expanded in a CF basis, and corrected

by use of known atomic data.

ICC Intraatomic Correlation Correction - An AIM calculation, in which basis set corrections due to rescaling are omitted.

IACC IntraAtomic Correlation Correction - An AIM calculation in which a FRS of CFs, formed from molecule optimized orthonormal PLOs, is used, and the corrections applied account only for the neglect of atomic correlation and relativity effects for each CF.

Two dimensionless units are used throughout. The atomic unit of length is the bohr; 1 bohr = 0.52917715 Å. The atomic unit of energy is the Hartree; 1 h = 27.212 eV = 627.51 Kcal/mole = 219475 cm^{-1}. The Hartree is a very large unit; chemical energy differences are usually measured in millihartrees (mh).

II. THE FORS MODEL AND ITS APPLICATION TO DIATOMIC MOLECULES

A. Introduction

The model of the Full Optimized Reaction Space (FORS) was developed to do justice to the desiderata for a theory of chemical reactions listed in Chapter I. This model was first introduced by Ruedenberg and Sundberg (1976) and has subsequently been applied to a number of reactions by Sundberg (1975), Cheung, Sundberg and Ruedenberg (1979), Dombek (1977), Feller (1979), Johnson and Schmidt (1981), and Feller, Schmidt and Ruedenberg (1982). More recently, the concept has been adopted by Roos et al. (1980) and Siegbahn et al. (1980) under the name "complete active space".

The model bears some relationship to certain previous approaches. One is that taken several years ago by Schaefer and Harris (1968) in their calculations using the full valence space generated from fixed minimal basis set Slater-type atomic orbitals. Others are represented by various MCSCF models with restricted configuration selection such as the Optimized Valence Configurations (OVC) approach of Wahl and Das (1977), the Hartree-Fock plus Proper Dissociation (HF+PD) model of Lie and Clementi (1974), the Generalized Valence Bond (GVB) method of Bobrowicz and Goddard (1977), or the separated pair model of Silver, Mehler and Ruedenberg (1970). Very similar in spirit to the present approach are the GVB-CI calculations of Hunt et al. (1972) and Dunning et al. (1976), and the Valence CI calculation of Kirby-Docken and Liu (1977). However, before the advent of the FORS model, there had been no attempt to combine consistently the concept of the full configuration

space with the principal of orbital optimization in this space and to explore systematically the implications of such a framework.

The unambiguous formulation of orbital and configurational spaces that are sufficiently general and unbiased to lead to quantitative potential energy surfaces is the major goal of the FORS model. This goal is obtained with the help of the minimal basis set concept.

B. Dominance of the Atomic Minimal Basis Set in Molecular Wavefunctions: Effective Molecular Bases

Experimental and theoretical chemists alike routinely use the concept of "formal minimal basis sets" of atomic orbitals when interpreting or anticipating the results of quantitative molecular calculations in an intuitive manner. By contrast, minimal basis sets usually do not play a prominent role in the actual execution of most current ab initio calculations, since extended AO bases have proven to be essential for obtaining quantitative results. Bardo and Ruedenberg (1974) as well as Feller and Ruedenberg (1979) have shown, however, that, in fact, minimal basis sets also dominate the AO expansions of accurate molecular ab initio wavefunctions. These authors calculated very accurate SCF wavefunctions in terms of large (up to 16s,8p,2d) uncontracted bases of Gaussian primitives and then determined "molecular-adapted" contractions of the same primitives in such a manner that SCF calculations in terms of these contracted bases would optimally reproduce the uncontracted calculations.

The aforementioned investigations led to the following conclusions regarding efficient and quantitatively reliable contracted AO bases:

(i) On each atom, there exists a small number of dominant contracted AOs whose occupation numbers exceed those of the rest by an order of magnitude at least. The number of these dominant contracted AOs on any atom is always equal to the number of minimal basis set AOs of that atom;

(ii) Only a small number of additional contracted AOs (one, two or three, depending upon the desired accuracy) is required to recover the results of the uncontracted calculations;

(iii) A large overlap (usually in excess of 0.9) exists between the orbital space spanned by the dominant molecular-adapted contractions mentioned under (i) and the orbital space spanned by the minimal basis set SCF AOs of the free atom;

(iv) The orbital space spanned by the dominant molecule-adapted contracted AOs mentioned under (i) and (ii) is practically identical with that obtained when the few lowest-lying virtual SCF AOs of the free atom are added to the minimal basis set free-atom SCF AOs mentioned under (iii);

(v) For the purpose of spanning the orbital space required to express molecular wavefunctions, the low-lying virtual free-atom SCF AOs can be replaced by the most diffuse Gaussian primitives. This is so because the lowest virtual free-atom AOs in a given basis of primitives are nearly identical with those AOs which are obtained by Schmidt orthogonalizing the most diffuse primitives to all occupied free-atom SCF AOs.

Another consideration is the choice of primitive atomic orbitals (PAOs) to be used in constructing the contracted basis set. Molecular calculations are usually performed with functions having a Gaussian radial dependence, as suggested by Boys (1950). A simple and convenient choice for the Gaussian exponents is the even-tempered criterion of Ruedenberg, Raffenetti and Bardo (1973), namely $\zeta_K = \alpha\beta^K$. Raffenetti (1975) has presented a few such Gaussian PAO bases. Feller and Ruedenberg (1979) and Schmidt and Ruedenberg (1979) have shown how the parameters α and β vary as the size of the primitive set is changed. Finally, Schmidt and Ruedenberg (1979) have presented Gaussian bases for hydrogen to argon, inclusive. These PAO sets are given for small to very large expansion lengths.

The ready availability of such PAO bases and the previously detailed conclusions regarding efficient contraction schemes suggests the following "recipe" for constructing basis sets to be used in molecular calculations:

PAOs: A set of even-tempered Gaussian primitives which approaches the atomic Hartree-Fock limit to the desired level of accuracy is chosen. These PAOs are contracted to two types of quantitative basis orbitals (QBOs).

Principal QBOs: The MBS of the occupied SCF-AOs of the free atoms. For hydrogen, the free-atom AO can be chosen with a scale factor of 1.2, provided that the hydrogen is not dissociated during the reaction of interest, in which case it should be unscaled.

Secondary QBOs: Single Gaussian primitives. They are of three kinds.

(i) The most diffuse PAOs that occur in the principal QBOs, as discussed above;

(ii) Polarization functions, i.e., PAOs with ℓ quantum numbers higher than those of the principal QBOs;

(iii) Very diffuse PAOs with the same ℓ quantum numbers as the principal QBOs, but whose orbital exponent is chosen as the even-tempered parameter α for atoms possessing appreciable negative charge in the molecule or α/β for Rydberg states.

Such extended basis sets were first suggested by Raffenetti (1973), and hence will be termed Raffenetti-type bases. Molecular calculations can be easily and routinely performed with such bases. They are at least as effective as any other type of contracted AO basis and, moreover, are easily constructed for any atom, for any arbitrary number of primitives and for any size of the contracted basis. They have the further advantage of explicitly containing the dominant QBOs, together with the less important QBO, which function as modifiers to the dominant QBOs to account for the molecular environment. Thereby, the rigorous AO bases remain no longer purely numerical devices, but become useful vehicles also for physically meaningful analyses of molecular wavefunctions - a goal that often seems to have been abandoned in commonly used ab initio bases (Dunning and Hay, 1977).

Such analyses confirm again the dominant role played by the principal QBOs and the reason for this dominance is readily understood in terms of the variation principle. It is the fact that any occupation

of the minimal-basis-set free-atom SCF AOs, even if only partial and indirect (i.e., through being building blocks of MOs), will yield substantially lower energy contributions than similar occupations of any other ("promoted") AOs that could be formulated. This conclusion remains also valid when going beyond the Hartree-Fock approximation and, consequently, it is to be expected that the electronic rearrangements occurring during chemical reactions can be essentially (i.e., in zeroth order) described as redistributions of electronic populations among the principal valence shell QBOs while preserving the overall dominance of this subset of basis orbitals. The surprisingly good results of the early calculations with fixed minimal-basis-sets of Slater-type AOs (Schaefer and Harris, 1968) support this inference.

Unless otherwise specified, the molecular basis sets used for the calculations described in this dissertation are of the Raffenetti-type described above, taken from Schmidt and Ruedenberg (1979). The following type of notation will be used to specify the basis: (14s,7p,2d/3s,2p,2d). The portion before the slash gives the number of Gaussian primitives (PAOs) used on a given atom, the part following the slash gives the number of contracted functions (QBOs) formed from these primitives according to the "recipe" given above.

C. The Full Optimized Reaction Space (FORS) Model

From the aforementioned "persistent dominance" of the principal quantitative basis orbitals, the following fundamental conjecture, which forms the basis for the concept of the "Full Optimized Reaction Space (FORS)", may be inferred:

The number of dominant configuration-generating molecular orbitals is at most equal to the number of formal minimal-basis-set atomic orbitals in the molecule.

Denoting this number by n, it is further postulated that meaningful descriptions of molecular rearrangements are obtained by expressing the electronic wavefunction in terms of n configuration-generating MOs (CGOs). No assumptions are made as regards the form of the CGOs and as regards the selection of configurations, except that the CGOs corresponding to closed shells remain doubly occupied in all CSFs. The CGOs are expected to be determined as expansions in terms of extended basis sets of principal and secondary QBOs by means of an MCSCF procedure.

The most general full reaction space (FRS) is the full valence space. Here it is only assumed that the CGOs describing inner shells remain doubly occupied in all configurations and all remaining CGOs are considered as "reactive orbitals". The FRS is then spanned by all possible configurations resulting from all possible couplings between the reaction orbitals and using all possible spin couplings. It is apparent that the inclusion of all valence configurations allows for a description of any possible electronic rearrangements associated with arbitrary atomic rearrangements, including dissociations and atomizations, without introducing any bias due to preconceptions or nescience. Under certain conditions, there may exist valid reasons for anticipating that certain CGOs can be chosen in such a manner, e.g., as lone pairs or bond orbitals, that they remain essentially closed, i.e., doubly occupied, during a specific reactive arrangement. In such cases, the number of

persistent closed shells can be extended to include these CGOs and this leads to a reduced full reaction space. In any event, the n CGOs of the system are divided into n_C closed core CGOs and n_R reactive CGOs, and the FRS is defined as the configuration space which is spanned by all configurations which keep the n_C core CGOs doubly occupied and couple the n_R reactive CGOs in all possible ways, compatible with the overall space and spin symmetry.

The configuration space which results when, for a specific molecular geometry, the CGOs of the full reaction space (FRS) are MCSCF optimized, is called the full optimized reaction space (FORS). This optimization in an adequately large orbital variation space is essential. Those orbitals in the orbital variation space which are orthogonal to the optimal CGOs are called virtual or unoccupied FORS MOs.

From the described construction, it is apparent that the CGOs of an FRS are not unique. This is so because the FRS is invariant with respect to the following two kinds of unitary orbital transformations:

(i) transformations among the n_C closed CGOs;

(ii) transformations among the n_R reactive CGOs.

The former invariance is entirely analogous to that known for Hartree-Fock wavefunctions and, in fact, leaves each configuration by itself invariant. The latter invariance gives rise to a unitary transformation between the configurations constructed from the "old" CGOs and those constructed from the "new" CGOs. This flexibility in the choice of the CGOs is a valuable asset of the FRS, since it offers the possibility of "adapting" the CGOs to various computational or interpretative purposes, a feature which will be examined in this and the following chapter.

D. Molecular Wavefunctions and the ALIS Programs

For the molecular calculations reported here, the wavefunction is chosen as a superposition of configurations

$$\psi^{SM} = \sum_{Kt} C_{Kt} \Phi^{SM}_{Kt}$$

where each configuration Φ^{SM}_{Kt} is chosen to be normalized spin-adapted antisymmetrized product (SAAP) of CGOs. A SAAP is an N electron wavefunction of the form

$$\Phi^{SM}_{Kt}(\text{space,spin}) = N_K A\{U_K(\text{space})\theta^{SM}_t(\text{spin})\}$$

where θ^{SM}_t is a spin eigenfunction, S and M being the eigenvalues of S^2 and S_z; U_K is a product of CGOs; $A = (N!)^{-1/2} \Sigma_p (-1)^P P$ is the conventional antisymmetrizer, and $N_K = 2^{-\pi(K)/2}$ with $\pi(K)$ being the number of doubly occupied CGOs in U_K. Properties of SAAPs are described by Ruedenberg (1971), Salmon, Cheung and Ruedenberg (1972), and Ruedenberg and Poshusta (1972). In cases of sufficiently high spatial symmetry, certain fixed combinations of SAAPs can be chosen as configuration state functions (CSFs). The molecular wavefunction is obtained by a multiconfiguration self-consistent-field (MCSCF) calculation, that is the energy functional is minimized with respect to the superposition coefficients C_{Kt} and the expansion of the CGOs in the QBO basis. Occasionally, only the superposition coefficients are optimized, with the CGOs held invariant; this type of wavefunction is known as a configuration interaction (CI) wavefunction. In the special case that only one configuration is used, so that only the QBO expansion of the CGOs need be

optimized, the wavefunction is referred to as the self-consistent-field (SCF) type.

The molecular calculations detailed here were all performed using the ALIS (Ames Laboratory-Iowa State) computer program system. The ALIS system contains a version of Raffenetti's (1973) BIGGMOLI program for generation of integrals over generally contracted QBOs; an integral transformation program by Elbert (1978), and an MCSCF program based on the Brillouin-Levy-Berthier theorem described by Ruedenberg, Cheung and Elbert (1979). The ALIS system is a very powerful tool for modern quantum mechanical investigation of reaction pathways of small molecules, and is now in worldwide usage. A number of other programs were used to draw orbital contour plots, obtain localized orbitals, or compute molecular properties.

E. Examples: Diatomics with Small Full Reaction Spaces

As a first example, consider the $^1\Sigma_g^+$ ground state of Li_2. The formal minimal basis set consists of the 1s, 2s, 2px, 2py, 2pz AOs on both atoms so that $n_C = 2$ and $n_R = 8$. In symmetry-adapted form, the CGOs are:

closed CGOs: $1\sigma_g$, $1\sigma_u$;

reaction CGOs: $2\sigma_g$, $3\sigma_g$, $2\sigma_u$, $3\sigma_u$, πx_u, πx_g, πy_u, πy_g .

From them, ten SAAPs can be constructed which possess $^1\Sigma_g^+$ symmetry. They are

$$|2\sigma_g^2\rangle,\ |3\sigma_g^2\rangle,\ |2\sigma_g 3\sigma_g\rangle,\ |\pi x_u^2\rangle,\ |\pi y_u^2\rangle\ ,$$

$$|2\sigma_u^2\rangle,\ |3\sigma_u^2\rangle,\ |2\sigma_u 3\sigma_u\rangle,\ |\pi x_g^2\rangle,\ |\pi y_g^2\rangle\ ,$$

where $|\phi^2\rangle$ and $|\phi\psi\rangle$ are defined by

$$|\phi^2\rangle = 2^{-3/2}\, A\{1\sigma_g^2 1\sigma_u^2 \phi^2 [(\alpha\beta - \beta\alpha)/\sqrt{2}]^3\} \quad .$$

$$|\phi\psi\rangle = 2^{-1}\, A\{1\sigma_g^2 1\sigma_u^2 \phi\psi [(\alpha\beta - \beta\alpha)/\sqrt{2}]^3\} \quad .$$

For $^1\Sigma_g^+$ symmetry, the πx^2 and πy^2 configurations must occur with equal coefficients so that they can be combined to the two configuration state functions (CSFs)

$$|\pi_u^2\rangle = \{|\pi x_u^2\rangle + |\pi y_u^2\rangle\}/\sqrt{2} \quad ,$$

$$|\pi_g^2\rangle = \{|\pi x_g^2\rangle + |\pi y_g^2\rangle\}/\sqrt{2} \quad .$$

Furthermore, because of the invariance of the FORS space, it is possible to choose the natural orbitals of the resulting wavefunction as CGOs. Since there are only two open shell electrons, transformation to these "natural reaction orbitals (NROs)" will annihilate the contributions of $|2\sigma_g 3\sigma_g\rangle$ and $|2\sigma_u 3\sigma_u\rangle$ to this wavefunction. Since the CGOs are going to be MCSCF-optimized, these two off-diagonal SAAPs can, therefore, be omitted from the start and, thus, the FORS space is actually only six dimensional.

For infinite separation, the wavefunction assumes the asymptotic form $\Psi = \{|2\sigma_g^2\rangle - |2\sigma_u^2\rangle\}/\sqrt{2}$. Here the invariance of the FRS can be used to introduce the "localized reaction orbitals (LROs)"

$$1\sigma_A = (1\sigma_g + 1\sigma_u)/\sqrt{2} \qquad 1\sigma_B = (1\sigma_g - 1\sigma_u)/\sqrt{2}$$

$$2\sigma_A = (2\sigma_g + 2\sigma_u)/\sqrt{2} \qquad 2\sigma_B = (2\sigma_g - 2\sigma_u)/\sqrt{2} \quad .$$

Using them as CGOs, one can transform the asymptotic wavefunction into the form

$$\Psi = 2^{-1} A\{1\sigma_A^2\ 1\sigma_B^2\ 2\sigma_A\ 2\sigma_B[(\alpha\beta - \beta\alpha)/\sqrt{2}]^3\}$$

$$= [2^{-1} A\{(1\sigma_A\alpha)(1\sigma_A\beta)(2\sigma_A\alpha)(1\sigma_B\alpha)(1\sigma_B\beta)(2\sigma_B\beta)\}$$

$$- 2^{-1} A\{(1\sigma_A\alpha)(1\sigma_A\beta)(2\sigma_A\beta)(1\sigma_B\alpha)(1\sigma_B\beta)(2\sigma_B\alpha)\}]/\sqrt{2} \quad .$$

At infinity, the localized CGOs become the atomic 1s and 2s SCF AOs, i.e., $1\sigma_A$, $1\sigma_B$, $2\sigma_A$, $2\sigma_B$ become identical with the principal QBOs $1s_A$, $2s_A$, $1s_B$, $2s_B$, and it is readily seen that the energy of Ψ becomes the sum of the atomic SCF energies. This is not always the case, however; an atom may dissociate into a two configuration wavefunction of the type $C_1 s^2 p^n + C_2 s^0 p^{n+2}$, if $n \leq 4$. This is a common situation for atoms such as beryllium, boron, or carbon.

An MCSCF optimization at the experimental bond length of 5.07 bohr, using a (12s,3p,1d/6s,3p,1d) Raffenetti-type QBO basis in this FORS space yielded an energy of -14.9006 h (h = one Hartree), can be compared with the SCF result of -14.8712 h and an SCF energy for two lithium atoms of -14.8651 h. Thus, this calculation improves the binding energy from the SCF value of 0.17 eV to the FORS value of 0.97 eV, within 1.8 Kcal/mole of the experimental value of 1.07 eV. Moreover, the FORS calculation connects Li_2 and 2Li by a good dissociation curve.

As a second example, consider the $^3\Sigma^-$ ground state of NH, imidogen. Here the minimal basis consists of the 1s, 2s, 2px, 2py, and 2pz orbitals of nitrogen, and the 1s orbital of hydrogen. Only the nitrogen 1s is a core orbital so that, in $C_{\infty v}$ notation, the CGOs are:

closed CGOs: 1σ ;

reaction CGOs: 2σ, 3σ, 4σ, $1\pi_x$, $1\pi_y$.

From these, nine $^3\Sigma^-$ SAAPs containing eight electrons, can be constructed:

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 2\sigma^2\ 4\sigma^2\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 2\sigma^2\ 3\sigma^1\ 4\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_3$$

$$1\sigma^2\ 3\sigma^2\ 4\sigma^2\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 2\sigma^2\ 3\sigma^1\ 4\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 3\sigma^2\ 2\sigma^1\ 4\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 3\sigma^2\ 2\sigma^1\ 4\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_3$$

$$1\sigma^2\ 4\sigma^2\ 2\sigma^1\ 3\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_1$$

$$1\sigma^2\ 4\sigma^2\ 2\sigma^1\ 3\sigma^1\ 1\pi_x^1\ 1\pi_y^1\ \theta_3 \quad .$$

The first SAAP is the SCF configuration around R_e, the first three SAAPs are required for proper dissociation to the ^4S and ^2S SCF ground states of nitrogen and hydrogen. Note that the three space orbital products

having four unpaired electrons are not combined with the overall triplet Serber spin function θ_2, which singlet couples the second pair of singly occupied orbitals to $^1\Delta$ symmetry, causing these SAAPs to possess an overall symmetry of $^3\Delta$.

Direct MCSCF optimization with a (14s,7p,2d/5s,3p,2d) QBO basis on nitrogen and an unscaled (6s,2p/3s,2p) QBO basis on hydrogen to obtain the FORS wavefunction at a variety of internuclear distances were performed. The potential curves are shown in Fig. 2.1, along with the SCF potential. It is apparent that the FORS curve, unlike the SCF curve, treats the dissociation reaction NH → N + H correctly, at least qualitatively. Selected spectroscopic constants, obtained via Dunham analysis with a program by Valtazanos (1980) are given in Table 2.1. The FORS calculation recovers 40% of the error in the SCF bond energy (D_e), but is still in error by 24.9 Kcal/mole.

As a final example of a diatomic with a small full reaction space, consider F_2. F_2 has the same CGOs as Li_2. Where Li_2 has two valence shell electrons, F_2 lacks two electrons of possessing a completely filled valence shell. Exploiting this fact by using a "hole" notation, the full valence space SAAPs for F_2 are:

$$
\begin{array}{lll}
|3\sigma_u^{-2}\rangle & |\pi x_g^{-2}\rangle & |2\sigma_u^{-1}\ 3\sigma_u^{-1}\rangle \\
|3\sigma_g^{-2}\rangle & |\pi y_g^{-2}\rangle & |2\sigma_g^{-1}\ 3\sigma_g^{-1}\rangle \quad , \\
|2\sigma_u^{-2}\rangle & |\pi x_u^{-2}\rangle & \\
|2\sigma_g^{-2}\rangle & |\pi y_u^{-2}\rangle &
\end{array}
$$

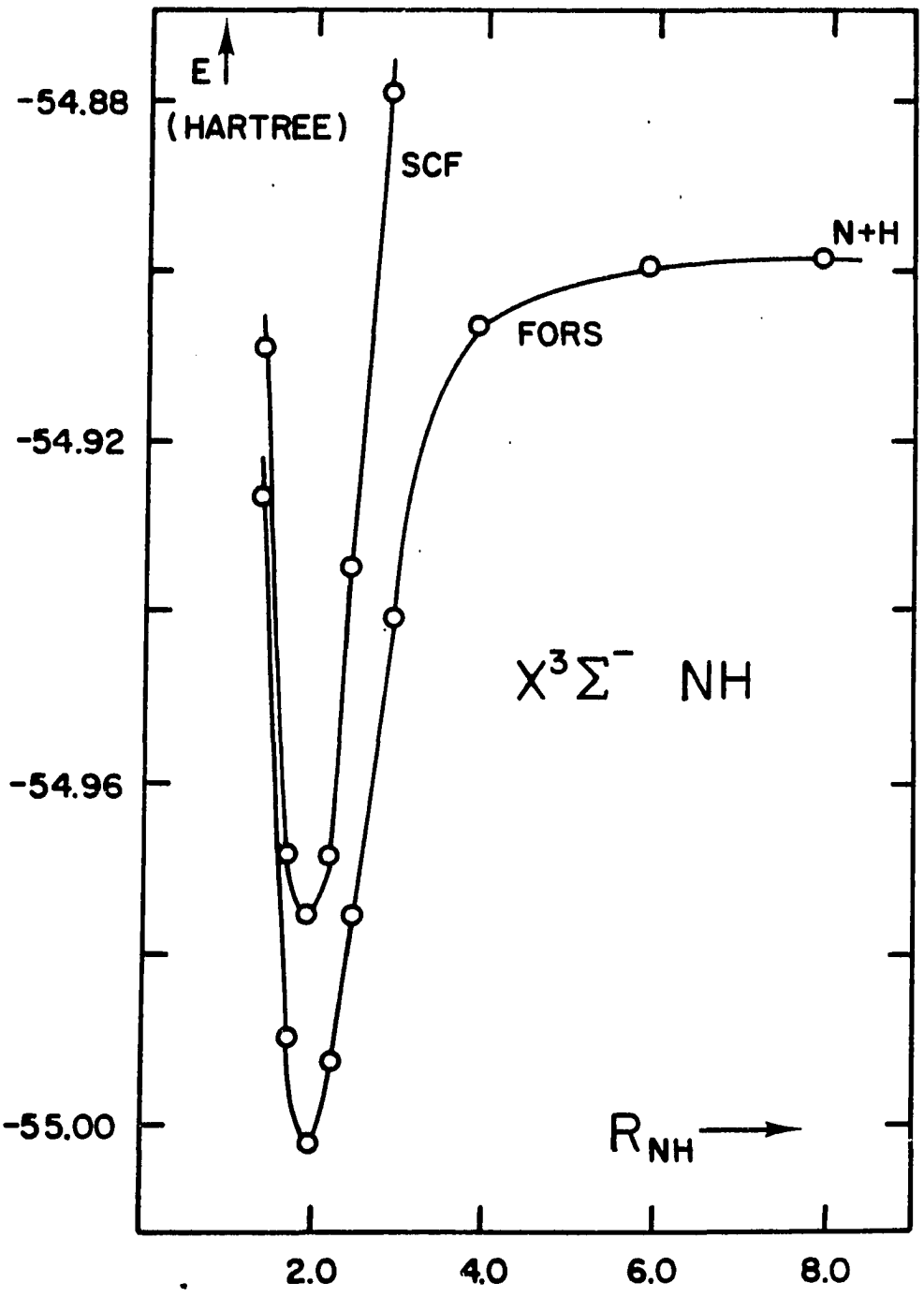


Figure 2.1. Calculated potentials for the ground state of imidogen

Table 2.1. Spectroscopic constants for $x^3\Sigma^-$ NH

	SCF	FORS	Experiment[a]
R_e (Å)	1.058	1.067	1.036
B_e (cm^{-1})	16.0	15.7	16.7
D_e (eV)	2.06	2.77	3.85 ± 0.10[b]
ω_e (cm^{-1})	3180	3090	3282
$\omega_e\chi_e$ (cm^{-1})	---	85	78.3
α_e (cm^{-1})	---	0.57	0.649

[a]Huber and Herzberg (1979).

[b]Piper (1979).

where the notation indicates which two electrons are to be deleted from the completely filled KKLL shell $1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 2\sigma_u^2 3\sigma_g^2 3\sigma_u^2 \pi_u^4 \pi_g^4$. The first configuration is the SCF configuration near R_e; the first two are required for proper dissociation into two SCF fluorine atoms. Just as for Li_2, the two SAAPs $|2\sigma_g^{-1} 3\sigma_g^{-1}>$ and $|2\sigma_u^{-1} 3\sigma_u^{-1}>$ are eliminated by transformation to natural orbitals, and the four π hole SAAPs combine to form two $^1\Sigma_g^+$ CSFs, so that the full valence space of F_2 consists of 6 CSFs.

F_2 is a rather famous example of the failings of SCF theory. Beyond the normal circumstance that the diatomic SCF function does not dissociate into a product of the two atomic SCF functions, SCF theory fails to predict F_2 as a bound molecule, as first noted by Wahl (1964). Using a (14s,7p,2d/4s,3p,2d) QBO basis at the experimental equilibrium distance, the SCF energy is -198.764060 h, corresponding to a bond energy of -1.40 eV, i.e., F_2 lies above 2F by 1.40 eV. MCSCF optimization of the FRS SAAPs yields an energy of -198.844328 h, representing a bond energy of 0.78 eV, and thus the satisfying conclusion that F_2 is a stable species. The FORS wavefunction removes 71% of the binding energy error of the SCF function, compared to the experimental value of 1.65 eV. Moreover, the FORS wavefunction of F_2 smoothly dissociates to two SCF F atoms. However, the calculated FORS bond energy for F_2 is in error by 20.1 Kcal/mole.

F. Methodologies for Large Full Reaction Spaces

The results for the three singly sigma bonded diatomics discussed in Section II.E indicate the FORS model to be of some usefulness for studying chemical reactions. When attempting to undertake such an

investigation, however, one is immediately confronted with many more configurations than for the simple systems thus far considered.

For diatomic molecules with s-p valence shells, the FRS generated by distributing in all possible ways all valence electrons in all CGOs which can be formed from all minimal basis set valence atomic orbitals can contain several hundred SAAPs. For triatomics, this full valence space can contain several thousand SAAPs; for ethylene, the full valence space consists of 29,248 SAAPs. It is clearly possible to conceive of full reaction spaces whose dimensions exceed the limitations of any MCSCF procedure, because there is a limit to the number of configurations that can be handled by any configuration interaction method, and even more severe limits exist in this respect for MCSCF procedures. It is, therefore, an important problem to establish reliable rules for identifying in advance as many noncontributing configurations as possible.

One method for significantly reducing the number of SAAPs has already been presented in Section II.C. A reduced FRS is obtained by demanding that certain bonds and/or lone pairs which intuitively seem to be uninvolved in a particular reaction remain doubly occupied. For example, Cheung et al. (1979) were able to study the breaking of the double bond in ethylene with a reduced FRS containing just eight rather than nearly 30,000 SAAPs by the simple expedient of requiring the four CH bond orbitals to remain doubly occupied.

In general, it is not always possible to achieve a reduction in the full reaction space to only a dozen or so SAAPs. More often than not,

interesting chemical problems require FRSs containing hundreds or thousands of SAAPs, even after all uninvolved lone pairs and bonds have been held to double occupancy. The MCSCF procedure used here can optimize only tens of SAAPs, which precludes the direct optimization of any FRS containing hundreds or thousands of SAAPs.

1. Optimization in a large FRS

Although it is not currently feasible to directly MCSCF optimize a FRS containing hundreds or thousands of SAAPs, this number of SAAPs is well within the reach of any configuration interaction (CI) procedure. Dombek (1977) has detailed a method that combines CI calculations in the FRS with a circumstance that eliminates the need for MCSCF optimization in the entire FRS. It happens that the shapes of the optimal CGOs depend only upon a relatively small number of leading configurations in the FRS. This is not to say that the remaining configurations are irrelevant for the energy calculation. Typically, at least 30% of all configurations in the FRS are required to recover the energy of the full CI calculation to within 0.1 millihartree. However, as a rule, less than 3% will be needed to determine the optimal molecular orbitals. To be precise: the energy obtained from the CI calculation in the entire FRS generated by the best possible MOs (i.e., the true FORS MOs) will differ by less than a millihartree from the energy obtained by a CI calculation in the entire FRS generated from those orbitals which are determined by means of an MCSCF calculation using only the leading 3% of all configurations. In practice, this almost always means less than twenty configurations are required for any MCSCF calculation.

Thus, MCSCF calculations are possible for very large reaction spaces, if it is possible to ascertain the dominant configurations. How is it possible to identify those before knowing the wavefunction which is yet to be determined? The answer is: by the macro-iterative approach of Dombek (1977).

Assume that one has an approximation to the optimal set of CGOs of a large full reaction space. Methods for obtaining such initial CGOs are discussed in the following subsections. These CGOs can be improved by the following two-step macro-iteration:

(i) a) A CI calculation is performed in the FRS spanned by these CGOs;

b) The symmetry-adapted natural orbitals of the CI wavefunction are determined;

c) The CI wavefunction is expressed in terms of NO-based configurations (i.e., configurations generated from the natural orbitals);

(ii) a) The leading NO-based configurations are selected from the expansion established under (i) c);

b) An MCSCF calculation is performed using these leading configurations to get improved CGOs.

The number and selection of the leading NO-based configurations for the MCSCF step depends on the coefficients of the NO-based CI expansion. Appropriately this number should be increased from one macro-iteration to another corresponding to the increasing refinement of the CGOs. These two-step macro-iterations are repeated until self-consistency is achieved

according to the following three convergence criteria:

(i) Stability of the CI energy in the full reaction space;

(ii) Stability of the NO-based CI expansions in the full reaction space as regards the order of the NO-based configurations and the values of the coefficients of these configurations;

(iii) Lack of change in the CGOs resulting from the MCSCF step.

Normally, convergence is attained after three macro-iterations involving two MCSCF steps, the first typically involving between five and ten configurations and the second involving between ten and twenty configurations. The MCSCF energies are clearly unsuited for monitoring the macro-iterative convergence.

The transformations to natural orbitals are necessary in order to be consistent in performing step (ii) a) of the macro-iterations and in applying the convergence criterion (ii). This is because, in general, the CGOs resulting from any one MCSCF step are not identical with the natural orbitals from that MCSCF wavefunction and, moreover, the number of leading configurations normally varies from one macro-iteration to another. Because of the invariance of the full reaction space under orbital transformation, convergence cannot be monitored for the coefficients of the CI expansions when the configurations are generated directly with the CGOs from the MCSCF step.

Orbital optimizations can also be simplified in other ways. There is no molecule (except H_2) in which it is not possible to divide the CGOs in groups such that there is only little coupling between the orbital optimizations in different groups. The efficiency of the MCSCF

process is substantially increased if, at first, orbital optimizations are performed separately in each group, freezing the orbitals in the other groups. Usually the final total MCSCF calculation will then converge in two iterations. Similarly, the efficiency of the MCSCF process is increased if orbital optimizations are performed first for a small basis of QBOs, to obtain excellent starting CGOs for the MCSCF calculation with a very large number of basis orbitals. A wise optimization strategy not only yields substantial savings in computer usage, but also furnishes valuable insights into the interactions that exist between various orbitals and configurations in a molecule.

It must be reemphasized that, even though the described MCSCF steps are very adequate to determine the optimal MOs, the energy must be calculated to the accuracy of the full reaction space in order to obtain reliably changes on energy surfaces for chemical reactions. Experience has shown that there exists no justifiable selection of a small number of configurations, such as 3% of the FRS, to represent a system without energy bias. Moreover, the leading configurations usually vary significantly along the reaction path. Retaining the entire FRS in the CI steps of the above macro-iterative procedure ensures that the leading configurations will always be found at any given point on the molecular potential energy surface.

2. Obtaining initial CGOs

The macro-iterative procedure of Dombek (1977) just outlined assumes that a reasonably accurate initial choice of CGOs is available. Practically speaking, poor convergence of Dombek's method is almost

always attributable to the lack of adequate starting orbitals. Consequently, one should make every effort to obtain the best possible starting CGOs.

In cases where a preliminary SCF calculation is feasible, a quite general procedure suggested by Sundberg and applied by Dombek (1977) can be applied. The occupied SCF MOs are taken as a zeroth order approximation for some of the FRS CGOs. However, since in general, there exist more CGOs in the FRS than there are occupied SCF MOs, additional zeroth order CGOs which lie in the virtual SCF space must be found. The lowest SCF virtual MOs are usually poor choices for these additional CGOs, due to their great spatial diffuseness. Instead, the necessary additional CGOs are obtained as follows. A few additional configurations are chosen which, for intuitive reasons, can be expected to contribute substantial contributions (typically double excitations providing left-right correlations in bonding MOs), just enough so that every CGO occurs in at least one configuration. Then, an MCSCF calculation is performed for a wavefunction consisting of the SCF configuration plus the newly chosen configurations, with the occupied SCF MOs frozen, i.e., only the additional CGOs in the virtual SCF space are optimized. This "frozen orbital" preliminary MCSCF calculation yields the desired full set of initial CGOs. A straight CI calculation with these initial CGOs will furnish the approximate wavefunction in the FRS from which the natural orbitals can be determined and the macro-iterative optimization in the FRS begun.

This technique of an SCF calculation followed by a "frozen orbital" MCSCF to generate starting CGOs for the macro-iterative optimization of the FRS is completely general, and was routinely applied in the calculations described below. The method assumes nothing about the structure of the FRS; it may be applied to any reduced FRS as readily as to a full valence reaction space.

It is worth noting that it is not necessary for the initial SCF wavefunction, on which the just described procedures are based, to have proper spatial and spin symmetry, even though the MOs must. This is so because the determination of molecular orbitals is the only objective of these preliminary calculations. Thus, any SCF calculation which converges is adequate, not necessarily one on the electronic state whose FORS function is sought. For example, a closed shell calculation may be performed in lieu of one with two unpaired electrons. It may be helpful to carry out the SCF calculation on the singly positive ion as this tends to generate more compact virtual orbitals.

An effective means of determining better virtual orbitals for initiating the "frozen orbital" MCSCF is to generate the modified virtual orbitals (MVOs) of Bauschlicher (1980). MVOs are obtained by deleting about five or so electrons from the converged SCF orbitals, and diagonalizing within the virtual orbital subspace the Fock operator corresponding to this highly positive ion. The lowest virtuals are "drawn in" to the region of the molecule inhabited by the valence MOs whose electrons were deleted. Thus, the expense of one SCF iteration generates MVOs which can save one or more iterations during a subsequent MCSCF

run. As an MCSCF cycle can easily be ten times longer than an SCF cycle, the savings are apparent.

Convergence difficulties are a common occurrence in MCSCF calculations. An inappropriate initial guess for the weakly occupied correlating orbitals can preclude convergence. Another common difficulty in optimizing a reduced FRS is the scrambling of bonds and lone pairs in the delocalized canonical SCF MOs. Localizing the canonical MOs, say, into a bond orbital which is to be correlated and a bond and a lone pair which are to remain doubly occupied could substantially improve convergence over letting this localization occur during the MCSCF cycles. Poor convergence can also be attributed to poor selection of configurations rather than poor starting orbitals. A wavefunction that contains mainly singly excited configurations usually converges slowly; the addition of a few sensible double excitations can effect convergence in only a few iterations. A final source of difficulty is an ill-conceived full reaction space. For example, a single orbital cannot be expected to correlate two different bonds. If this is attempted, the MCSCF cycles will oscillate wildly; in one cycle one bond is correlated, in the next cycle the other bond. Adding a second correlating orbital to the CGO set, thus expanding the FRS, will immediately bring convergence.

A FORS function for any particular molecule can perhaps be obtained readily, or perhaps with some difficulty. When encountered, poor convergence can be met by a cleverer choice of initial orbitals, wiser selection of configurations, or a more appropriate FRS. MCSCF calculations in general, and FORS calculations in particular, require a

theorist willing to think through convergence problems instead of just allowing many more expensive MCSCF cycles.

3. An alternative choice of initial CGOs: the atomic MBS

In introducing the FORS model in Section II.C, the number of molecular orbitals, or CGOs, in the full valence space was chosen equal to the number of formal minimal basis set (MBS) AOs on all atoms in the molecule. This is so because the low lying molecular orbitals (higher occupied and lower unoccupied) from which effective configurations are to be formed are expected to be bonding and antibonding combinations of these MBS AOs. As a consequence, one would expect the full valence space spanned by the configurations generated from the minimal basis set of these free-atom SCF AOs to be quite similar to the Full Optimized Reaction Space of the molecule. Indeed, this turns out to be true and it proves to be useful because, for a minimal basis set, the orbital reaction space is identical with the orbital variation space and, hence, the optimal wavefunction in this approximate space can be obtained without orbital optimization by a straight CI calculation. Because of the approximate nature of such a preliminary analysis, it is desirable to choose as an orbital basis CGOs which lead to as clear a separation into strongly and weakly contributing configurations as possible. The weights of the dominant configurations will then be predicted with higher relative accuracy than those of the weak configurations. In general, natural orbitals are known to yield configuration expansions which have such a "compact" character.

Thus, excellent approximations to the FORS NOs and to the weights of the configurations based on FORS NOs can be obtained in advance of the macro-iterative optimization of the FRS by performing a straight CI calculation in the FRS generated by the principal QBOs, i.e., the atomic MBS, then finding the natural reaction orbitals of the resulting wavefunction, and finally repeating the CI calculation in terms of FRS configurations constructed from these natural orbitals.

When performing this kind of calculation, henceforth designated a MBS FRS CI calculation, it is necessary first to construct, from the principal QBOs, a basis of orthogonal approximate CGOs for the closed core space and for the reactive CGO space. Similarly, an orthogonal orbital basis for the corresponding virtual space can be constructed from the secondary QBOs. This is accomplished as follows. First, the inner shell QBOs are mutually orthogonalized to form the approximate closed core basis. Then the principal valence QBOs are Schmidt orthogonalized to the closed core orbitals just formed and then mutually orthogonalized to form the approximate basis of the reactive orbital space. Finally, the secondary QBOs are Schmidt orthogonalized to both the closed core and the reactive orbital space and then mutually orthogonalized to form a basis for the virtual space. (The latter is done with a view to subsequent calculations with the extended basis.)

This means of preparing starting CGOs for optimization of a full valence space-type wavefunction is easily implemented when Raffenetti-type bases are used. These bases, as described in Section II.B above, contain explicit core, valence, and virtual QBOs. Linear combinations

of these QBOs that are symmetry adapted to the molecular point group are easily formed. These linear combinations serve to span the initial core, valence, and virtual CGO spaces, and must be orthogonalized as described above. A generalization that produces starting CGOs for a reduced FRS is presented elsewhere by Ruedenberg et al. (1982).

The major advantage of using the MBS FRS CI natural orbitals as starting CGOs is that the MBS CI expansion, after transformation to natural orbitals, is quite similar to that of the optimized wavefunction in the FRS, expressed in terms of the optimized natural orbitals. Thus, before commencing the macro-iterative optimizations, the dominant configurations in the final optimized wavefunction are known, and can be immediately used in the MCSCF orbital optimization steps. In addition, the MBS FRS CI natural orbitals are adequate initial guesses for all CGOs, since as many MBS AOs as there are CGOs in the full valence space, by definition. In contrast, the starting CGOs obtained form an SCF calculation in the extended basis set give good approximations to the heavily occupied final CGOs, but rather poor approximations to the weakly occupied final CGOs. Lastly, since the procedure is described here for the most general FRS possible, the full valence space, there is no possibility of mistakenly working in an overly restricted FRS. Thus, the MBS FRS CI approach addresses the three major sources of convergence difficulties mentioned at the end of Section II.F.2.

4. Illustration of the MBS FRS CI alternative: F_2

As an example of using natural orbitals from a CI calculation in the FRS spanned by taking the MBS atomic orbitals as CGOs, consider

again the F_2 molecule. The optimized full valence space wavefunction for F_2 was described in Section II.E above. Here the use of MBS FRS CI natural orbitals is illustrated for F_2. Denoting atomic orbitals on the left or right fluorine as L or R, the core CGOs are taken as $1\sigma_g = 1s_L + 1s_R$ and $1\sigma_u = 1s_L - 1s_R$. The valence shell CGOs are taken as $2\sigma_g = 2s_L + 2s_R$, $3\sigma_g = 2pz_L + 2pz_R$, $2\sigma_u = 2s_L - 2s_R$, $3\sigma_u = 2pz_L - 2pz_R$, $\pi x_u = px_L + px_R$, and $\pi x_g = px_L - px_R$. The πy CGOs are analogous to the πxs. A CI in the FRS with these CGOs, after orthogonalization, yields natural orbitals, for which the $2\sigma_g/3\sigma_g$ and $2\sigma_u/3\sigma_u$ natural orbitals are mixtures of the originally pure 2s or 2pz $2\sigma_g/3\sigma_g$ and $2\sigma_u/3\sigma_u$ CGOs. A second CI in the FRS with these natural orbitals yields the expansion shown in Table 2.2 along with the optimal natural orbital based expansion.

The economically generated MBS FRS CI expansion is strikingly similar to the optimized FORS expansion. Clearly, the MBS FRS CI calculation correctly predicts those configurations which dominate the optimal wavefunction in the FRS. Besides generating such an accurate CI expansion in the FRS, the MBS FRS CI natural orbitals are excellent starting orbitals for the direct MCSCF optimization of this full valence space; only four MCSCF iterations are required for convergence.

The first configuration in Table 2.2 is the SCF wavefunction which corresponds to the standard bonding picture in terms of symmetry orbitals, resulting in a single sigma bond. From the mixing coefficients in the second column, it is seen to have a weight of 93.1% in the FORS wavefunction. There is one secondary configuration, with a weight of

Table 2.2. Mixing of configurations generated from natural orbitals in F_2 at the equilibrium distance

Configuration State Functions		Configuration Mixing Coefficients	
Hole Description	Excitation from SCF	FORS	MBS FRS CI
$3\sigma_u^{-2}$	SCF	0.965127	0.963282
$3\sigma_g^{-2}$	$3\sigma_g^2 \rightarrow 3\sigma_u^2$	-0.256618	-0.262606
$2\sigma_u^{-2}$	$2\sigma_u^2 \rightarrow 3\sigma_u^2$	-0.025857	-0.016167
$2\sigma_g^{-2}$	$2\sigma_g^2 \rightarrow 3\sigma_u^2$	-0.015070	-0.014130
$2\sigma_u^{-1}3\sigma_u^{-1}$	$2\sigma_u \rightarrow 3\sigma_u$	0	0
$2\sigma_g^{-1}3\sigma_g^{-1}$	$2\sigma_g 3\sigma_g \rightarrow 3\sigma_u^2$	0	0
$(\pi x_g^{-2} + \pi y_g^{-2})/\sqrt{2}$	$\pi_g^2 \rightarrow 3\sigma_u^2$	-0.040769	-0.046774
$(\pi x_u^{-2} + \pi y_u^{-2})/\sqrt{2}$	$\pi_u^2 \rightarrow 3\sigma_u^2$	-0.010898	-0.006944

6.6%, which corresponds to the left-right correlation in the sigma bond. The remaining configurations, adding up to a total of 0.3%, describe the other possible degeneracy-type correlations among the principal QBOs. As noted in Section II.E, the coefficients of the two configurations with singly occupied orbitals are annihilated by transformation to natural orbitals, and the π "holes" must be combined as shown to obtain $^1\Sigma_g^+$ CSFs.

The effectiveness of the MBS FRS CI procedure derives from the fact that the vector space spanned by the minimal basis set of the free atom AOs has a very high overlap with the optimal orbital reaction space spanned by the FORS CGOs. Quantitative demonstration of this overlap will be given in Chapter III. In view of such very large overlap values, it is justified to raise the question whether the minimal basis set may not be adequate by itself, and whether the MCSCF calculation in the orbital variation space of the extended basis may, therefore, not be superfluous. The answer is provided by the results of the calculations of F_2 given in Table 2.3. It is seen that the relatively slight improvements in the FORS MOs due to the extended basis makes a difference of about 23 Kcal/mole in the energy, raises the correlation energy recovery from 39% to 71% and is essential for obtaining binding of F_2 in the FRS approach. Thus, the energetic effect is substantial and of practical consequence. This is presumably so because the orbital basis of the principal QBOs, although being "good" for the description of the molecule, is even better for the separated atoms, for which it is, in fact, optimal. Thus, there results a bias in favor of the separated

Table 2.3. Energies for various F_2 calculations

Type of Calculation	Total Energy (h)	Binding Energy (eV)
SCF = FORS of two F atoms	-198.815576	
SCF of F_2 molecule	-198.764060	-1.40
MBS FRS CI of F_2	-198.807886	-0.21
FORS MCSCF of F_2	-198.844328	+0.78
Experimental		+1.65

atoms which leads to a binding energy that is too low and, moreover, distorts the energy curve so severely that the equilibrium distance is displaced outward by about 0.5 bohr for most diatomics. This explanation is in agreement with the observation that minimal basis set CI calculations with Slater-type AOs, which are relatively poor for the separated atoms as well as for the molecule, usually yield better energy differences between the molecule and the separated atoms than do minimal basis set CI calculations with free-atom SCF AOs (Schaefer and Harris, 1968).

As illustrated for F_2, the MBS FRS CI wavefunction typically does not recover as much correlation energy as optimization in the FRS does, and the potential surfaces from such calculations are not particularly parallel to the exact surfaces. These two reasons necessitate optimization in the FRS, for which the MBS FRS CI wavefunction is an excellent starting point.

G. Examples: Diatomics with Large Full Reaction Spaces

As examples of optimization of a large full reaction space, consider two triply bonded diatomics, CO and N_2, at their experimental equilibrium distances using rather extensive basis sets. Both molecules have ten configuration generating orbitals (CGOs) arising from the 1s, 2s, and 2p orbitals of the constituent atoms and ten reactive electrons. The only real difference in the structure of the FRS for these two molecules is the absence of an inversion center in CO. CO and N_2 are chosen to illustrate the various techniques described in Section II.F to optimize in a large FRS.

In CO, the closed core consists of the (1σ) and (2σ) CGOs describing the inner shells; the reactive CGOs are (3σ), (4σ), (5σ), (6σ); (1πx), (2πx); (1πy), (2πy). Using ten reactive electrons, 316 SAAPs can be constructed which can be combined to 176 CSFs of $^1\Sigma^+$ symmetry. This is the full valence space which is chosen as the FRS. Upon dissociation in the FRS, one obtains the SCF ground state of oxygen and a two-configuration wavefunction of the form $[a(s^2p^2) + b(s^0p^4)]$ for the ground state of carbon.

Starting orbitals are obtained by the MBS FRS CI method described in Section II.F.3. The entire calculation converges in 2 1/2 macro-iterations involving two MCSCF calculations. Pertinent details of the optimization are presented in Table 2.4. In the first macro-iteration, the FRS CI calculation is based on the principal QBOs. Casting the resulting CI expansion in terms of configurations based on natural orbitals yields the type and order of configuration from which the six SAAPs for the subsequent MCSCF calculations are deduced. In this MCSCF calculation, the two core CGOs corresponding to the inner shells are kept frozen. In the second macro-iteration, the FRS CI is based on the CGOs from the first MCSCF calculation and, from the natural-orbital-based reexpansion of this CI wavefunction, the 19 leading SAAPs are taken for the second MCSCF calculation in which <u>all</u> CGOs (core and reactive) are optimized. It may be noted that the CI energy of the first macro-iteration (in the FRS generated by the principal QBOs) is about 3 eV poorer than the SCF energy, which again illustrates the energetic importance of the secondary QBOs. Nonetheless and in agreement with the

Table 2.4. FORS calculations for CO and N_2[a]

Property	Results of FORS Calculation SCF	FORS	EXP
Carbon Monoxide[b]			
Energy (h) of C	-37.6883	-37.7057	
Energy (h) of O	-74.8084	-74.8084	
Energy (h) of CO	-112.7829	-112.9144	
Binding energy (eV)	7.79	10.89	11.225
Dipole moment (Debye)	+0.26	-0.30	-0.12
Nitrogen Molecule[c]			
Energy (h) of N	-54.4004	-54.4004	
Energy (h) of N_2	-108.9853	-109.1337	
Binding energy (eV)	5.01	9.06	9.90

Macro-iteration	Energies During Iterative Process Step 1: FRS-CI	Step 2: MCSCF
Carbon Monoxide		
1	-112.6858	-112.8676 (6 SAAPs/5 iterations)
2	-112.9122	-112.9089 (19 SAAPs/2 iterations)
3	-112.9144	
Nitrogen Molecule		
1		-108.9853 (SCF)
2	-109.0847	-109.0780 (8 SAAPs/4 iterations)
3	-109.1180	-109.1222 (13 SAAPs/7 iterations)
4	-109.1334	-109.1281 (19 SAAPs/3 iterations)
5	-109.1337	-109.1292 (23 SAAPs/1 iteration)
6	-109.1337	

[a]Geometries: R_{CO} = 2.13195, R_{N_2} = 2.068 bohr.

[b]CO basis: (14s,7p,2d/4s,3p,2d) d exponents SCF optimized in CO from Feller and Ruedenberg (1979), 0.2567 and 1.001 for C, 0.1813 and 1.0643 for O.

[c]N_2 basis: (14s,7p,2d/5s,3p,2d) d exponents assumed as 0.2 and 1.0.

discussion in Section II.F.3, the expansion of this CI wavefunction in terms of natural-orbital-based configurations is quite similar to the analogous expansion of the final converged CI calculation, and its natural orbitals are excellent starting CGOs for the first MCSCF calculation.

In N_2, the closed core is spanned by the CGOs $1\sigma_g$ and $1\sigma_u$, while the reactive CGOs are $2\sigma_g$, $3\sigma_g$, $2\sigma_u$, $3\sigma_u$; πx_u, πx_g; πy_u, πy_g. The number of possible SAAPs is 176 and they can be combined to 95 CSFs of $^1\Sigma_g^+$ symmetry. The FORS wavefunction dissociates into the SCF description of the ground states of the two nitrogen atoms.

The details of the N_2 calculation are also shown in Table 2.4. For this optimization, the starting CGOs were obtained from the SCF plus "frozen orbital" MCSCF procedure described in Section II.F.2. A preliminary SCF calculation was followed by a preliminary MCSCF calculation which was based on eight configurations selected from the NO-based expansion of the CI calculation in the FRS spanned by the occupied and the lowest three virtual SCF MOs. In this preliminary MCSCF calculation, four CGOs were frozen and six were optimized. Subsequently, three macro-iterations were carried through using 13, 19 and 23 configurations, respectively. The final MCSCF calculation was unnecessary; its object was merely to demonstrate that the use of 23 configurations would not alter the results obtained by using 19 configurations: the CGOs and the FRS-CI energy remained indeed unchanged. In fact, the FRS-CI energy obtained with the CGOs from the 13 configuration MCSCF calculation is already within 0.3 millihartree of the final FORS energy. By comparing

the CO and N_2 calculations, it can be seen that the MBS FRS CI approach gets more quickly into effective macro-iterations and eliminates the need for orbital optimization prior to the first FRS-CI calculation.

From the results for bond strengths of CO and N_2 given in Table 2.4, it is apparent that the effectiveness of the FORS model is qualitatively similar in these two isoelectronic molecules. It is interesting to note, however, that this consistency depends on the consistent application of the FORS concept, which demands the earlier mentioned two-configuration wavefunction for the carbon atom. If instead the one-configuration SCF approximation is used for the carbon atom, as is mistakenly done by Kirby-Docken and Liu (1977), then the value 11.32 eV is obtained for the dissociation energy of CO, which is larger than the experimental value.

The binding energy is not the only molecular property which is improved by the FORS wavefunction. Table 2.4 also contains dipole moments for CO. The dipole moment of CO is an oft-quoted failure of SCF theory, as first noted by Huo (1965). SCF theory predicts a small positive dipole for CO, corresponding to C^+O^-. The FORS calculation correctly predicts a small negative dipole, meaning a negative carbon. The discrepancy between the FORS and experimental dipole is due to the rapid change in the dipole moment with bond length. It is this large dipole derivative that makes CO an intense infrared absorber. The FORS moment is calculated at R_e, and the experimental moment is averaged over the molecular vibration. Billingsley and Krauss (1974) found vibrational averaging of their OVC function for CO slightly reduced the magnitude

of their small negative moment, the same effect would be expected for the FORS function.

H. Diatomics as a Test of the FORS Model

The crucial test for a chemical theory of molecular potential energy surfaces is how well it predicts reaction energies. How accurate are the calculated exo- or endoergicities? How accurate are the predicted barrier heights? Diatomics serve as excellent benchmarks for such a theory. An accurate experimental dissociation energy, D_e, that is the reaction energy for $AB \rightleftarrows A + B$, is usually available for comparison. Only one molecular calculation, at the known equilibrium position, R_e, need be performed to obtain the D_e predicted by the theory. The atomic bases can be chosen large enough to approach the single particle basis set limit, so that errors in the calculated D_e are attributable to failings in the theory rather than the basis used to execute the theory. Finally, diatomics are a stringent test of a theory's ability to calculate the variations in electron correlation at different locations on the potential surface. Where the atoms typically have one or more unpaired electrons in a high spin state, the diatoms are low spin systems possessing newly paired electron bonds.

Accordingly, diatomics are presented here as a test of the accuracy of the FORS model. Tables 2.5 through 2.9 contain a summary of SCF and FORS calculations on all homonuclear and several important heteronuclear diatomics of the second period. These tables draw, in part, on unpublished calculations of Brenda Lam and David Feller, as indicated in Table 2.5.

Table 2.5. Synopsis of diatomic calculations

Diatomic Molecule	Electronic State	Bond Distance[a] (bohr)	QBO[b] Basis Set	Calculation Performed by
		Homonuclear:		
Li_2	$x^2\Sigma_g^+$	5.07	(12s,3p,1d/6s,3p,1d)[c]	Schmidt
B_2	$x^3\Sigma_g^-$	3.0047	(14s,7p,2d/4s,3p,2d)[d]	Lam
C_2	$x^1\Sigma_g^+$	2.579	(14s,7p,2d/4s,3p,2d)[e]	Lam
N_2	$x^1\Sigma_g^+$	2.068	(14s,7p,2d/5s,3p,2d)[e]	Schmidt
O_2	$x^3\Sigma_g^-$	2.2817	(14s,7p,2d/4s,3p,2d)[f]	Schmidt
F_2	$x^1\Sigma_g^+$	2.68	(14s,7p,2d/4s,3p,2d)[f]	Schmidt
		Heteronuclear:		
CN	$x^2\Sigma^+$	2.2144	(14s,7p,2d/4s,3p,2d)[g]	Schmidt
BO	$x^2\Sigma^+$	2.2762	(14s,7p,2d/4s,3p,2d)[h]	Lam
CO	$x^1\Sigma^+$	2.132	(14s,7p,2d/4s,3p,2d)[i]	Schmidt
NO	$x^2\pi$	2.1747	(14s,7p,2d/5s,3p,2d)[j]	Feller

[a]Experimental equilibrium distance from Huber and Herzberg (1979).

[b]Atomic basis from Schmidt and Ruedenberg (1979). Polarization as indicated.

[c]ζ_P = 0.0678, 0.264, 1.03; ζ_D = 0.275 SCF optimized in Li_2.

[d]ζ_D = 0.145, 0.913 assumed values.

[e]ζ_D = 0.2, 1.0 assumed values.

[f]ζ_D = 0.5, 1.6 assumed values.

[g]C: ζ_D = 0.3, 1.05 N: ζ_D = 0.32, 1.12 assumed values.

[h]B: ζ_D = 0.145, 0.913 O: ζ_D = 0.15, 1.1 assumed values.

[i]C: ζ_D = 0.2567, 1.001 O: ζ_D = 0.1813, 1.0643 SCF optimized in CO (Feller and Ruedenberg, 1979).

[j]N: ζ_D = 0.257, 1.0 O: ζ_D = 0.1809, 1.062 assumed values.

The calculations described in this section are all on the ground electronic states of the diatoms. To increase the number of molecules surveyed, the calculations were only performed at the experimental equilibrium distances. The minimum of the SCF potential often occurs at a slightly smaller internuclear distance; FORS wavefunctions almost inevitably yield somewhat elongated bonds. In either case, the calculated dissociation energy would be increased only slightly by geometry optimization. (These assertions are not true for the inflexibly contracted MBS FRS CI calculations, see the discussion in Section II.F.4.)

All diatomic calculations described here are the same in spirit as the Li_2, F_2, N_2, and CO calculations discussed above. Two core CGOs represent the atomic inner shells; eight reactive CGOs arising from the atomic 2s and 2p shells are used to construct configurations correlating each valence electron. The number of SAAPs required for these full valence space calculations is shown in Table 2.7. The FORS space of each diatomic considered dissociates to atomic ground states.

Accurate chemical predictions require the error in calculated bond energies to be no more than kT, which at room temperature is 0.9 millihartree, 0.025 eV, or 0.6 Kcal/mole. Roughly speaking, the error in a calculated energy derives from two sources, namely from using a finite number of configurations, and secondly, in the orbitals due to their expansion in a finite basis. To ensure a valid test of a configurational selection scheme such as the FORS model, the molecular basis sets used ideally should approach the limit of completeness to within kT to eliminate this second type of error.

The molecular basis sets, shown in Table 2.5, consist of atomic basis sets involving extensive numbers of Gaussian primitives, contracted flexibly (triple zeta or better), with at least two sets of polarization functions. The atomic bases used approach the atomic limit to within about kT; their errors vary monotonically from 0.1 millihartree for Li to 1.6 millihartree for F (Schmidt and Ruedenberg, 1979). The molecular basis sets, as shown in Table 2.6, have errors of about five to ten kT. This error is estimated as the difference between the molecular SCF functions calculated here and the results of SCF calculations using optimized extensive Slater (exponential) type orbital bases. The errors given in Table 2.6 are described as maxima because the possible small improvements in the atomic bases presumably carry over into the diatom and make no contribution to the calculated D_es. The remainder of the molecular error is from two causes, omission of an f polarization orbital on atoms other than Li, and the lack of optimization of the d polarization orbitals. This residual error does contribute to increased D_es, so that the estimated basis errors in Table 2.6 should not be counted against the FORS model.

The energies of the SCF and FORS wavefunctions are shown in Tables 2.6 and 2.7, respectively. These data are shown converted into dissociation energies in Table 2.8. It is apparent that the FORS dissociation energies are much superior to the SCF values. A convenient measure of the accuracy of a wavefunction is the percentage of "molecular extra correlation energy" recovered, that is, the percentage of the discrepancy between SCF and actual dissociation energies. In most cases, the FORS

Table 2.6. SCF calculations on diatomics

Diatomic Molecule	Molecules SCF Energy (hartree) (this work)	Atomic SCF[a] Energy (hartree) (this work)	Molecular SCF[b] Energy (hartree) (literature)	Maximum Basis Error[c] (millihartree)	Maximum Basis Error[c] (Kcal/mole)
		Homonuclear:			
Li_2	-14.8712	-7.4326	-14.8715[d]	0.3	0.2
B_2	-49.0874	-24.5289	-49.0909[d]	3.5	2.2
C_2	-75.4015	-37.6883	-75.4062[d]	4.7	2.9
N_2	-108.9853	-54.4004	-108.9928[d]	7.5	4.7
O_2	-149.6580	-74.8084	-149.6659[d]	7.9	5.0
F_2	-198.7641	-99.4078	-198.7701[d]	6.0	3.8
		Heteronuclear:			
CN	-92.2192	---[e]	-92.2232[f]	4.0	2.5
BO	-99.5566	---[e]	-99.5555[g]	-1.1[h]	-0.7[h]
CO	-112.7829	---[e]	-112.7892[i]	6.3	4.0
NO	-129.2894	---[e]	-129.2953[j]	5.9	3.7

[a]Calculated using a 14s,7p basis (12s for Li) from Schmidt and Ruedenberg (1979).

[b]Lowest published energies, using extensive spdf STO basis sets.

[c]Estimated as column 2 minus column 4. See text for discussion.

[d]Cade and Wahl (1974).

[e]See entries for the two corresponding homonuclear diatoms.

[f]Green (1972a).

[g]Cade and Huo (1975).

[h]The data here indicate the B0 wavefunction in reference g is about 5 millihartree above the Hartree-Fock limit.

[i]Green (1970).

[j]Green (1972b).

Table 2.7. FORS calculations on diatomics

Diatomic Molecule	Number of[a] SAAPs	Molecular FORS Energy (hartree)	Atomic FORS Energy (hartree)
		homonuclear:	
Li_2	8	-14.9006	-7.4326
B_2	136	-49.2171	-24.5601[b]
C_2	264	-75.6369	-37.7056[b]
N_2	176	-109.1337	-54.4004
O_2	44	-149.7631	-74.8084
F_2	8	-198.8444	-99.4078
		heteronuclear:	
CN	616	-92.3708	---[c]
BO	616	-99.6780	---[c]
CO	316	-112.9144	---[c]
NO	252	-129.4055	---[c]

[a]This is the dimension of the full valence space, i.e., all electrons correlated by all configurations built from all valence orbitals. The number of $C_{\infty v}$ adapted CSFs is in general smaller.

[b]The ground states of these atoms have two CSF FORS functions of the form $as^2p^n + bs^0p^{n+2}$. The FORS function of the other atoms is the SCF function.

[c]See the entries for the two corresponding homonuclear diatoms.

Table 2.8. Diatomic dissociation energies

Diatomic Molecule	D_e(SCF)[a] (eV)	D_e(FORS)[b] (eV)	D_e(experiment)[c] (eV)	% Correlation[d] Recovered	D_e^{FORS} Error[e] (Kcal/mole)
			homonuclear:		
Li	0.16	0.96	1.068	88	2.5
B_2	0.81	2.64	3.08	81	10.1
C_2	0.68	6.14	6.32	97	4.1
N_2	5.02	9.06	9.905	83	19.5
O_2	1.12	3.98	5.213	70	28.4
F_2	-1.40	0.78	1.658	71	20.2
			heteronuclear:		
CN	3.55	7.21	7.89	84	15.7
BO	5.97	8.42	8.40	101	-0.5
CO	7.79	10.89	11.226	90	7.7
NO	2.19	5.35	6.615	71	29.2

[a] It is traditional to define this quantity as the atomic SCF energies minus the molecular SCF energy. Note that the SCF atoms and molecule are not connected by a smooth potential curve.

[b] The difference between the FORS atoms and molecule. These are connected by a smooth curve.

[c] Calculated from data in Huber and Herzberg (1979) as $D_e = D_0^0 + 1/2\ \omega_e - 1/4\ \omega_e x_e + 1/8\ \omega_e y_e$.

[d] $100\% \times (D_e(\text{FORS}) - D_e(\text{SCF}))/(D_e(\text{experiment}) - D_e(\text{SCF}))$.

[e] D_e (experiment) - D_e(FORS).

wavefunction recovers 70-95% of this correlation error, which represents a very substantial improvement over the SCF value. On the other hand, the remaining error in the FORS bond energies is 5 to 30 Kcal/mole, which is several times larger than the 2 to 5 Kcal/mole error attributable to the basis set.

These residual errors in the FORS bond energies seem quite unsystematic, except for two general trends. Except for BO, all calculated FORS dissociation energies are too small. Also, as might be expected, the FORS model does relatively better when the number of valence electrons is smaller than the number of valence orbitals. In Li_2, which has one electron pair and seven orbitals to correlate it, 88% of the correlation error is recovered. In F_2, which has seven pairs and one correlating orbital, 71% is recovered. Oddly enough, the variation in the effectiveness of the FORS model for the isoelectronic pair CO and N_2 is not significantly different than for any other pair of diatoms.

Another important molecular property is the dipole moment. The calculated SCF and FORS dipoles of the heteronuclear molecules are compared to the experimental values in Table 2.9. In all cases, the FORS values are an improvement over the SCF values, and are quite close to the experimental values. As noted in Section II.G, failure to average the calculated dipole over the molecular vibration is probably the largest source of error in the FORS dipoles. For the ground vibrational level, the averaged dipole μ_o is obtained from the calculated dipole μ_e at R_e by the formula (Raymonda, Muenter, Klemperer, 1970)

Table 2.9. Diatomic dipole moments[a]

Diatomic Molecule	SCF	FORS	Experiment[b]
CN	2.30	1.62	1.45 (C^+N^-)
BO	3.00	2.34	- (B^+O^-)
CO	-0.26	0.30	0.122 (C^-O^+)
NO	-0.31	0.24	0.159 (N^-O^+)

[a]All values given in Debye units.

[b]Huber and Herzberg (1979).

$$\mu_o = \mu_e + \frac{1}{2}\left(\frac{\alpha_e}{B_e} + 3 \frac{B_e}{\omega_e} \right) R_e \left. \frac{d\mu}{dR} \right|_{R_e} + \ldots \quad .$$

This vibrational averaging cannot be done without additional information on the potential curve (ω_e,α_e) and the dipole derivative function ($d\mu/dR$).

The calculations reported here primarily serve to test the accuracy of the FORS model. However, the ultimate goal of a theory is to make predictions for unknown chemical properties. This is possible for BO, which has received little experimental or theoretical attention. It is clear from Table 2.6 that the Slater type orbital bases used in SCF calculations by Cade and Huo (1975) are not as well-optimized as for the other molecules. Except for BO, the FORS bond energies in Table 2.8 recover 70-95% of the correlation error, and some 84% for the isoelectronic CN. It seems unlikely that the FORS model is any better for BO than the other diatoms, and thus, that the experimental bond energy is in error. Assuming the FORS wavefunction for BO recovers 85 ± 10% of the correlation energy, the predicted bond energy for BO is 8.85 ± 0.3 eV. This prediction is somewhat larger than the recent thermochemical determination of 8.44 ± 0.12 eV by Uy and Drowart (1970). It should be noted that the bond energy of BO has never been well-determined experimentally; published values range from 7.4 to 9.2 eV. Finally, Table 2.9 gives the dipole for BO as 2.34D, probably accurate to ±0.3D, for which no experimental result is available.

The FORS model underestimates bond energies for diatomics by five to thirty Kcal/mole. This error is larger than can be attributed to the basis sets used, and is a factor of ten larger than desired for reliable chemical predictions. As noted, diatomics are an extreme case of changing correlation energy as the new bond or multiple bonds are formed. In cases where the number of electron pairs does not change (e.g., closed shell reactants → closed shell products), the changes in electron correlation are much less severe, and the performance of the FORS model should be correspondingly better. In such situations, the FORS model can be expected to yield chemical energy differences accurate to a few Kcal/mole.

III. ORBITAL ANALYSES OF FORS WAVEFUNCTIONS

A. Introduction

As shown in Chapter II, the Full Optimized Reaction Space model was introduced to obtain wavefunctions which describe electronic rearrangements without bias along an entire chemical reaction pathway. A second major advantage is that the FORS wavefunctions are invariant under certain transformations of the occupied MOs. The usefulness of this invariance for chemical interpretation of the wavefunction is the subject of this chapter.

The FORS wavefunction is invariant against two types of transformations, namely an arbitrary orthogonal transformation of the core, or doubly occupied orbitals, and an arbitrary orthogonal transformation of the reactive, or partially occupied valence orbitals. There are, thus, an infinite number of orbitals from which the same FORS wavefunction can be constructed, and a number of these orbital choices afford valuable chemical insight into bonding and other aspects of electronic structure.

Within the closed core orbital space are localized orbitals or canonical orbitals. The choices for the reactive orbitals are more numerous, and rather more interesting. Natural reaction orbitals yield the most compact FORS function possible. A new projection scheme yields molecular orbitals that are almost entirely localized onto one or another atom of the molecule and, therefore, furnish an unambiguous atomic population analysis. Appropriate compromises between these localized orbitals and the delocalized natural orbitals can be found to illustrate

various features of the chemical binding. These latter sets are termed chemically adapted orbitals.

In the next four sections, the manner in which the invariance of the closed core and reactive valence orbitals can be exploited is considered in detail. In the following section, some of the various orbital types are illustrated for some of the diatomic wavefunctions given in Chapter II, and chemical interpretation given. In the final section, the FORS orbitals are used to reexamine and confirm the definition of the FORS wavefunction.

B. Invariance of the Core Orbital Space

The invariance of each SAAP against orthogonal transformations among the closed core CGOs derives from the fact that these orbitals all have identical occupation numbers (i.e., 2), which leads to a core degeneracy of the first-order density matrix so that the natural orbitals for the core are not unique.

1. Canonical core CGOs

As in the analogous case of the restricted HF-SCF approximation, there exists a single Fock operator for all core CGOs and thus canonical core CGOs can be obtained by diagonalizing this core Fock operator F_C, given by

$$F_C = h + 2 \sum_j^C \{C_{jj} - 1/2\ K_{jj}\} + \sum_{k,\ell}^R p_{k\ell}\{C_{k\ell} - 1/2\ K_{k\ell}\} \qquad (3.1)$$

where the sums go over the core or reactive orbitals as indicated; h, C, and K are the usual one electron Hamiltonian, and two electron Coulomb

and exchange operators, and p is the bond order, or density matrix. A derivation of this operator is given by Das and Wahl (1966).

The formation of canonical core CGOs has the effect of separating the inner shell core MOs from those valence MOs which are included in the closed core. (These two types of MOs can get arbitrarily mixed by MCSCF optimizations.) The separation of inner shell and valence core MOs is important for conceptual interpretations as well as for computational refinements, such as following the FORS optimization with a CI calculation including excitations from these higher energy "core" MOs.

2. Localized core CGOs

On the other hand, the closed shell CGOs may be localized according to standard techniques, such as the Edmiston-Ruedenberg (1963, 1965, 1971) or the Foster-Boys (1960) methods. Such localization yields atomic inner shell MOs as well as those atomic lone pair MOs and localized bond MOs which are "unaffected" by the reaction and for this reason were included in the closed core, in order to reduce the dimension of the full reaction space.

In the case where the full valence space is used, the localized atomic inner shell MOs can also be obtained by the projection method to be defined in the next section.

C. Invariance of the Reactive Orbital Space

The invariance of the full reaction space (FRS) against transformations among the reactive CGOs derives from the fact that all possible configurations involving the CGOs are included in the FRS. An orthogonal

transformation among the reactive CGOs does not leave the individual configurations invariant, but it induces an orthogonal transformation among all configurations spanning the FRS. The FRS as an entity is thus invariant to arbitrary transformations of the reactive orbitals.

1. Natural reaction orbitals

A CGO set of particular interest is the one which furnishes the most compact and succinct characterization of the molecular wavefunctions, a goal which may be identified with the most compact orbital expansion of the first order density kernel. The CGOs which accomplish this objective are the natural reaction orbitals (NROs) (Löwdin, 1955), obtained by diagonalizing the first order density matrix in the space of the reactive CGOs. This diagonalization tends to extremize the eigenvalues of the density matrix, which are the orbital occupations. For closed shell molecules near their equilibrium geometry, this separates the natural orbitals into those which are nearly doubly occupied and those which are largely vacant. Molecules with unpaired electrons will also have NOs with occupancies near one.

It is this separation of the orbitals into dominant (approximately singly or doubly occupied) orbitals and secondary (nearly zero occupation) correlating orbitals that causes the FORS wavefunction expressed in the NO basis to be the most compact form. As noted in Chapter II, this compactness facilitates the choice of a small number of configurations to be used during the MCSCF orbital optimization steps.

For nuclear geometries away from the equilibrium positions of reactants or products, i.e., near transition states, the occupations can

assume any value from zero to two, and the NO based FORS expansions are then somewhat less compact than at the equilibrium positions. Obtaining the FORS wavefunction at a transition state is only marginally more difficult than at the equilibrium position (a few more configurations are needed in the MCSCF orbital optimization). This is in sharp contrast to experimental chemical methods, where fleeting intermediates are virtually inaccessible.

2. Localized reaction orbitals

The reactive valence orbitals, that is, those with occupancies less than two, may be localized according to any intrinsic criterion, such as the maximization of the sum of the orbital self energies (Edmiston and Ruedenberg (1963, 1965)) or minimizing the sum of the squares of orbital dipoles (Foster and Boys (1960)). Normally these localization criteria are applied to the doubly occupied orbitals of an SCF function, whose single configuration is invariant to such localization. Here the localization can be applied to any or all the reactive orbitals, which are not doubly occupied, since the total FORS wavefunction is invariant to this localization. The resulting localized orbitals are termed localized reaction orbitals (LROs).

It is apparent that localized reaction orbitals usually do not belong to irreducible representations of the molecular symmetry group. Therefore, the number of SAAPs constructed from LROs, which are required to span the full configurational reaction space, is usually larger than the number of SAAPs constructed from symmetry-adapted CGOs required to span the FRS. For a molecule of average symmetry (say C_{2v}), this can

increase the number of SAAPs in the FORS function by a factor of ten. Calculation of the FORS wavefunction based on LROs rather than symmetry-adapted orbitals such as the NROs requires correspondingly greater computer resources. The major usefulness of LROs is, therefore, interpretive rather than computational.

Fortunately, the bond order (first order density) matrix for the LROs can be obtained without recalculating the FORS wavefunction. Given the orthogonal transformation T from any initial symmetry adapted MOs ϕ to the LROs ψ

$$\psi = \phi T \quad , \tag{3.2}$$

the bond order matrix over the LROs is calculated by

$$\gamma^{\psi} = T^{+} \gamma^{\phi} T \quad . \tag{3.3}$$

A convenient choice of initial symmetry orbitals for the localization are the above mentioned NROs, for which γ^{ϕ} is diagonal. The diagonal elements of γ^{ψ} are the electron occupations of the LROs; the offdiagonal elements are related to bonding or antibonding effects between any pair of LROs.

Upon localizing the reactive CGOs of a FORS wavefunction, one nearly always obtains localized CGOs with the following characteristics:

(i) Each LRO is nearly completely localized on one atom, and the orthogonalizing tails on the neighboring atoms are extremely small; and

(ii) The number of LROs on any one atom is equal to the number of principal QBOs, i.e., of formal minimal basis set AOs on that

atom, which were included in determining the total number of reactive CGOs.

The latter characteristic is a consequence of, and confirmation for, the definition of the FORS wavefunction. The first characteristic means LROs are ideal vehicles for demonstrating atomic participation in molecular wavefunctions.

Consider the simple FORS wavefunction describing a single sigma bond. Upon MCSCF optimization of the two CGOs, one obtains a NRO with nearly double occupancy, corresponding to the sigma bond, and a just barely occupied NRO, which is a sigma antibond correlating the bond pair. Upon localization of these two NROs, one obtains a single orbital on each atom, directed toward the other atom. Each orbital will have an occupation of about one, indicating each atom contributes a single electron to the bond pair. Furthermore, the offdiagonal bond order matrix element will be a large positive number, indicating strong bonding between the atoms.

In general, the FORS wavefunction correlates more than one electron pair, so that after localization each atom in the molecule will possess more than one LRO. These LROs will generally not be readily interpreted as lone pairs, bond orbitals, etc. However, LROs having these interpretations can be found as transformations among the orbitals on any one atom preserve the localization onto that atom. A first step toward finding such LROs is to diagonalize the atomic submatrices of the bond order matrix. This diagonalization will produce on each atom three possible types of "atomic natural orbitals": nearly doubly occupied LROs

which are interpreted as lone pairs, nearly singly occupied orbitals which are involved in bonding, and occasionally, in the case of electron deficiency, nearly vacant LROs. Finally, when an atom participates in two or more bonds, it may well happen that the two or more singly occupied LROs obtained by the diagonalization of the interatomic density matrix do not point in the directions of the atoms to which it is bonded. This direction into bonds can be accomplished by further transformation of only the singly occupied orbitals on any one atom so as to maximize certain offdiagonal elements of the interatomic bond order matrix.

LROs which have been somehow localized onto atoms and then subjected to transformations on each atom to obtain lone pairs, and orbitals directed into bonds are termed directed localized reaction orbitals (DLROs). A beautiful illustration of the DLROs involved in the HNO-HON isomerization is given by Dombek (1977). DLROs generally look like hybrid atomic orbitals, for example, in methane, the DLROs would correspond to an s orbital on each hydrogen and four equivalent sp^3 hybrids on the central carbon. As such, the DLROs reveal how the constituent atoms participate in molecular bonding.

To maintain its orthogonality to the other LROs, each LRO possesses a small antibonding "tail" on adjacent atoms. Since these tails are small, it seems reasonable to assign the electron occupancy (diagonal element of the bond order matrix) of the LRO entirely to the atom on which the LRO principally resides. Summing over all LROs on an atom gives an unambiguous measure of the number of electrons populating that atom in the molecular wavefunction, a useful quantity in discussing charge

transfer. These population numbers are free of the limitations associated with the Mulliken (1955) population analysis, which requires the arbitrary partitioning of overlap populations between nonorthogonal atomic basis orbitals.

D. Projected Localized Orbitals

The localized core and reactive orbitals discussed above are obtained in two steps, namely localization onto the atoms, then transformations among the reactive orbitals on the same atom. The first step can be accomplished according to some criterion involving an intrinsic property of the orbitals, such as maximizing the self-energy sum in the Edmiston-Ruedenberg (1963, 1965) procedure. It is possible to achieve this initial localization onto atoms according to an extensive criterion; that the localized orbitals resemble as far as possible the basis orbitals located on that atom. The procedure serves to localize the core, reactive, and virtual MOs, requires only the overlap integrals for the atomic basis, and uses only seconds of computer time-orders of magnitude less than an energy localization.

The localization is accomplished by projection of the atomic basis orbitals onto the core, reactive, or virtual orbital subspaces of the FORS orbital space. The orbitals which result from the projection process are termed projected localized orbitals, and are of three types: projected localized core, reactive, or virtual orbitals (PLCOs, PLROs, PLVOs). The projection procedure is given here only for the case of the Full Valence Space type FORS wavefunction, where the reactive MOs are equal in number to the atomic valence shell orbitals. A possible

generalization to the situation where some valence MOs are incorporated into the core orbital subspace to achieve a FRS containing a reduced number of configurations is presented elsewhere (Ruedenberg, Schmidt and Gilbert, 1982).

Following optimization of the FRS, the molecular orbitals ϕ have a known expansion in the atomic orbital basis χ,

$$\phi = \chi C \quad . \tag{3.4}$$

Here ϕ and χ are row vectors, and the AOs have a known overlap matrix, while the MOs are an orthonormal basis,

$$S = \chi^{+}\chi \tag{3.5}$$

$$I = \phi^{+}\phi \quad . \tag{3.6}$$

For simplicity, it is assumed the MOs have no symmetry (C_1), and are ordered with the core MOs before the reactive MOs before the virtual MOs. The generalization to symmetry adapted MOs is presented below. The AOs must be placed in a corresponding order, those AOs which are chosen for projection onto the core MOs first, followed by the AOs to be projected onto the reactive and virtual MO subspaces. Equation (3.4) can be inverted, so that the AOs are expressed in the MO basis,

$$\chi = \phi D' \quad . \tag{3.7}$$

Matrix inversion is not necessary, D' is obtained by

$$\begin{aligned} D' &= ID' \\ &= \phi^{+}\phi D' \end{aligned}$$

$$= \phi^{+}\chi$$

$$= C^{+}\chi^{+}\chi$$

$$= C^{+}S \quad . \tag{3.8}$$

In Eq. (3.7), each AO χ_i is a linear combination of all MOs ϕ_k. A particular χ_i can be projected onto one of the core, reactive, or virtual MO subspaces by zeroing its coefficients D'_{ki} whenever the index k falls in the other two orbital subspaces. That is, the projections π of the AOs on the MOs are given by

$$\pi = \phi D \quad , \tag{3.9}$$

where, because of the initial ordering of the MOs and AOs, D is obtained from D' by zeroing all elements of D' outside the diagonal core, reactive and virtual blocks. These projections are neither orthogonal or normalized; their metric, or overlap matrix, is calculated by

$$\sigma = \pi^{+}\pi$$

$$= D^{+}\phi^{+}\phi D$$

$$= D^{+}ID$$

$$= D^{+}D \quad . \tag{3.10}$$

The set of orthonormal projected orbitals which most closely resembles these projections is obtained by symmetric orthogonalization (Löwdin, 1950). Let U be the orthogonal matrix which diagonalizes σ,

$$\lambda = U^{+}\sigma U \quad .$$

The matrix needed for the orthonormalization is

$$\sigma^{-1/2} = U\lambda^{-1/2}U^{+} \quad , \tag{3.11}$$

where the elements of $\lambda^{-1/2}$ are defined as

$$\lambda_{ij}^{-1/2} = \delta_{ij}/\sqrt{\lambda_i} \quad .$$

The final orthonormal projected orbitals are calculated from

$$p = \pi\sigma^{-1/2} \quad . \tag{3.12}$$

It is readily verified that these orbitals have a unit metric, and can be expressed in the original MO basis or the AO basis by

$$p = \phi T \tag{3.13}$$

$$p = \chi L \quad . \tag{3.14}$$

Substitution of Eqs. (3.9) and (3.4) into (3.12) shows T and L are calculated by

$$T = D\sigma^{-1/2} \tag{3.15}$$

$$L = CD\sigma^{-1/2} \quad . \tag{3.16}$$

Since p and ϕ are orthogonal orbital sets, the matrix T is an orthogonal transformation.

The occupied projected orbitals strongly resemble the original atomic orbitals which were projected. Since an AO is located on a single atom, its projection, which is of course an optimal molecular orbital capable of generating the original FORS wavefunction, is highly localized on that atom. Usually the overlap between one of the occupied

projected orbitals and the AO before projection is in excess of 0.9, so that the projected orbitals are indeed atom localized.

The projection scheme is most easily applied if the atomic orbital basis is of the Raffenetti type described in Chapter II. In this case, the single basis functions which represent the various atomic inner shells are projected onto the core MOs. The single basis functions representing the atomic valence orbitals are projected onto the reactive MOs, which are equal in number due to the choice of the full valence space type of FORS function. Lastly, all the secondary basis functions, such as diffuse and polarizing Gaussians, are projected onto the virtual MOs. The resulting PLOs strongly resemble the original s, p, d, etc. AOs, and hence could be termed s, p, d, etc. type orbitals, although they are molecular orbitals, and thus possess small orthogonalizing "tails" on adjacent atoms. A possible computational improvement to the FORS model, based on these unhybridized, strikingly atomic PLOs is presented in Chapter VI.

It is possible to apply the projection scheme when a segmented type basis is used. In this case, each atomic inner shell AO is a linear combination of the basis orbitals on that atom, and the valence AOs are different linear combinations of the same basis orbitals. These linear combinations are projected onto the core or valence MOs, as appropriate, to obtain the PLOs. The linear combinations of basis orbitals that are the orthogonal complement to the occupied AOs can be used to localize in the virtual space.

It is, of course, not necessary to project s, p, d, etc. type AOs. One could first form hybrid AOs from the atomic basis, and project these hybrids onto the FORS orbitals to obtain hybridized PLOs. The generalization of the projection scheme to the case of a reduced FRS given by Ruedenberg, Schmidt and Gilbert (1982) involves the construction of appropriate hybrids to be projected onto the reduced reactive orbital space.

As previously stated, the PLOs are quite similar to the AOs; their overlap is normally larger than 0.9. The major cause of this small difference is the symmetric orthonormalization in Eq. (3.12). If this orthonormalization is not performed, and instead the projections are just normalized according to

$$p = \pi N \quad , \quad N_{ij} = \delta_{ij}/\sqrt{\sigma_{ii}} \quad , \tag{3.17}$$

the resulting orbitals p will lack orthogonalizing components on adjacent atoms. In this case, Eqs. (3.15) and (3.16) are replaced by

$$T = DN \tag{3.18}$$

and

$$L = CDN \quad . \tag{3.19}$$

It should be noted that this T is not orthogonal, since it no longer relates two orthonormal bases. The absence of orthogonalizing tails on these nonorthogonal projections increases their atomic character (as measured by overlap with the AOs before projection). However, the difficulty of performing any type of calculation with a nonorthogonal

orbital basis renders these nonorthogonal projected orbitals useless from a computational standpoint.

Many molecules possess one or more symmetry elements, and the difficulty of the FORS optimization is alleviated substantially by performing it using symmetry adapted orbitals, which permits the symmetry classification of the individual configurations. Upon completion of the optimization, the molecular orbitals in each irreducible representation are either core, reactive, or virtual orbitals. The atomic basis functions are of course not symmetry adapted, but primitive symmetry orbitals can readily be formed from them by an orthogonal transformation U

$$\hat{\chi} = \chi U \quad . \tag{3.20}$$

These symmetry orbitals have the metric

$$\begin{aligned} \hat{S} &= \hat{\chi}^{+}\hat{\chi} \\ &= U^{+}\chi^{+}\chi U \\ &= U^{+}SU \quad , \end{aligned} \tag{3.21}$$

and the MOs can be expressed in the symmetry orbital basis by

$$\begin{aligned} \phi &= \chi C \\ &= \hat{\chi}U^{+}C \\ &= \hat{\chi}\hat{C} \quad , \end{aligned} \tag{3.22}$$

$$\hat{C} = U^{+}C \tag{3.23}$$

Both $\hat{S}$ and $\hat{C}$ are block diagonal along irrep boundaries.

The projection scheme is analogous to the case of no symmetry. The symmetry adapted AOs are expressed in terms of MOs as

$$\hat{\chi} = \phi\hat{D}' \tag{3.24}$$

$$\hat{D}' = \hat{C}^{+}\hat{S} \quad . \tag{3.25}$$

The projections are again given by zeroing elements of $\hat{D}'$ outside core, reactive, or virtual blocks, where this projection is now done in each block (irrep) of $\hat{D}'$. The resulting matrix $\hat{D}$ gives the unnormalized, nonorthogonal projections

$$\hat{\pi} = \phi\hat{D} \quad , \tag{3.26}$$

whose metric is

$$\hat{\sigma} = \hat{D}^{+}\hat{D} \quad . \tag{3.27}$$

Symmetrically orthonormal projections are obtained from

$$\hat{p} = \hat{\pi}\hat{\sigma}^{-1/2} \quad .$$

These symmetry adapted projections may now be localized onto the atoms by an orthogonal transformation W,

$$p = \hat{p}W \quad . \tag{3.28}$$

A possible choice for W is U^{+}, the inverse of the transformation from AOs to symmetry orbitals, although one could choose W differently, for example, to achieve a partial localization. The final PLOs are given by

$$p = \phi T \tag{3.29}$$

$$p = \chi L \tag{3.30}$$

where

$$T = \hat{D}\hat{\sigma}^{-1/2}W \tag{3.31}$$

$$L = C\hat{D}\hat{\sigma}^{-1/2}W \quad . \tag{3.32}$$

The use of symmetry to block diagonalize the matrices $\hat{C}$, $\hat{S}$, $\hat{D}$, $\hat{\sigma}$, etc. increases the speed of the projection process. Looping over the smaller diagonal blocks decreases the matrix multiplication and diagonalization time requirements, as well as storage needs.

Forming symmetry orbitals by an orthogonal transformation U, projecting the symmetry orbitals, and localizing with U^+ is equivalent to projecting without regard to symmetry. This is proven by showing the orbitals from the first method, $\hat{p}U^+$, are identical to those of the second, p. The two sets of projected orbitals are given by

$$\hat{p}U^+ = \hat{\chi}\hat{C}\hat{D}\hat{\sigma}^{-1/2}U^+ \tag{3.33}$$

$$p = \chi CD\sigma^{-1/2} \quad . \tag{3.34}$$

Before projection, Eq. (3.25) gives

$$\begin{aligned} \hat{D}' &= \hat{C}^+\hat{S} \\ &= (C^+U)(U^+SU) \\ &= (C^+S)U \\ &= D'U \quad , \end{aligned} \tag{3.35}$$

where Eqs. (3.23), (3.21) and (3.8) were used. This same relation holds after projection, as the symmetry adaptation U combines only equivalent

AOs, i.e., core with core AOs, valence with valence AO, and secondary with secondary AO, so

$$\hat{D} = DU \quad . \tag{3.36}$$

With this result,

$$\begin{aligned} \hat{\sigma} &= \hat{D}^{+}\hat{D} \\ &= U^{+}D^{+}DU \\ &= U^{+}\sigma U \quad , \end{aligned}$$

where Eqs. (3.27) and (3.10) were used. It can be shown that if two matrices are related by a similarity transformation, any power of these matrices are related by the same similarity transformation,

$$\hat{\sigma}^{m} = U^{+}\sigma^{m}U \quad , \tag{3.37}$$

in particular for m = -1/2. Substituting Eqs. (3.21), (3.23), (3.36) and (3.37) into (3.33)

$$\begin{aligned} \hat{p}U^{+} &= \hat{\chi}\hat{C}\hat{D}\hat{\sigma}^{-1/2}U^{+} \\ &= (\chi U)(U^{+}C)(DU)(U^{+}\sigma^{-1/2}U)U^{+} \\ &= \chi C D \sigma^{-1/2} \\ &= p \quad , \end{aligned}$$

according to (3.34). Q.E.D.

A very useful quantity in discussing orbitals and their significance in terms of occupation or contribution to bonding is the bond order matrix γ^{ϕ}, which represents the expansion of the first order reduced

density matrix in a particular orbital basis

$$\rho(r,r') = \sum_i \sum_j \gamma^{\phi}_{ij} \phi_i(r) \phi^*_j(r') \quad . \tag{3.38}$$

The first order density, i.e., the bond order matrix in a particular orbital basis can be calculated from the wavefunction obtained in the same orbital basis according to its formal definition

$$\rho(r,r') = N \int ds_1 \int d\tau_2 \ldots \int d\tau_N \, (rs_1,\tau_2, \ldots, \tau_N)\psi^*(r's_1,\tau_2, \ldots, \tau_N) .$$

Upon completion of the FORS optimization, the bond order matrix in the orbital basis used to construct the FORS wavefunction is, therefore, known. Changing the orbital basis by transformations within the core or reactive orbital spaces leaves ψ and thus ρ invariant, but the bond order matrix does change, and should thus be given a superscript label to indicate the orbital basis used to expand ρ. Upon localization, the bond order matrix in the PLO basis can be found without the computational cost of reexpressing the wavefunction in the low symmetry PLO basis. These PLOs are obtained by transforming the original MOs,

$$p = \phi T \quad ,$$

where T is given by equations such as (3.15) or (3.18). This can be inverted,

$$\phi = pT^{-1} \quad ,$$

with T^{-1} equal to T^+ if the PLOs are orthonormal, or else found by explicit matrix inversion if nonorthogonal PLOs are made. Substituting in Eq. (3.38) gives

$$\rho = \sum_i \sum_j \gamma^{\phi}_{ij} \left(\sum_k p_k T^{-1}_{ki} \right) \left(\sum_\ell p_\ell T^{-1}_{\ell j} \right)^*$$

$$= \sum_k \sum_\ell \left(\sum_i \sum_j T^{-1}_{ki} \gamma^{\phi}_{ij} T^{-1^+}_{j\ell} \right) p_k p^*_\ell \quad .$$

Of course, in the PLO basis

$$\rho = \sum_k \sum_\ell \gamma^{p}_{k\ell} p_k p^*_\ell \quad .$$

Comparing and using matrix notation,

$$\gamma^{p} = T^{-1} \gamma^{\phi} T^{-1^+} \quad . \tag{3.39}$$

The bond order matrix for a particular orbital basis is very helpful in interpretation of those orbitals. To repeat, the diagonal elements of the bond order matrix have the interpretation of electron occupations, while the sign and magnitude of offdiagonal elements are related to the strength of binding or antibinding interactions between orbitals.

Having obtained the bond order matrix over the PLOs, the procedure to obtain DLROs given in the previous section can be applied. The localization by projection can thus serve to nearly eliminate the computational requirements for obtaining DLROs, namely the initial localization onto atoms, which is quite time consuming for other localization methods.

E. Chemically Adapted MOs

The atom localized orbitals such as the unhybridized PLOs or the hybrid DLROs, and their bond order, or density matrix, serve to

characterize the participation of atoms in the molecular wavefunction. It is also desirable to find MOs which display the typical molecular characteristics of the electronic wavefunction. Such orbitals are called chemically adapted, and are generalizations of the energy localized MOs of an SCF wavefunction. The term chemically adapted is applied to any orbitals which clarify some aspect of the molecular bonding; thus, it may be possible or desirable to isolate several sets of chemically adapted MOs.

Chemical terminology for the description of electronic structure includes "two center bond", "resonance structure", "lone pair", "inner shell", "nonbonding orbital", "three center bond", "Lewis structure", etc. A general procedure to extract MOs from the FORS orbital space which correspond to these concepts is difficult to formulate, particularly for molecules with unusual electronic structure. Two different procedures are given below, however, any particular molecule may require some other treatment to best illustrate the electronic distribution in that molecule.

The discussion of chemical adaptation of the FORS MOs is presented here for a general FORS function, not necessarily a full valence space. Chemically adapted MOs for the closed shell orbital space are readily found by energy localization, as described in Section III.B.2. A guiding principle for the determination of chemically adapted reactive orbitals is maximal occupancy, that is, concentrating the electron distribution into as many nearly doubly occupied orbitals as possible. The natural reaction orbitals, found by diagonalizing the entire density matrix,

satisfy this criterion, but are unfortunately delocalized over the entire molecule. A second desirable quality of chemically adapted orbitals is that they should be localized into two center bonds, etc. Chemically adapted MOs thus respresent a compromise between highly localized atom adapted MOs and the extremized occupation of natural reaction orbitals. They should be as localized as possible, while possessing a nearly diagonal matrix. This is obviously not a unique definition, leaving room for maneuvering to determine the most clarifying possible chemical MOs.

One simple method for obtaining these MOs starts from the natural orbitals, which can usually be divided into three sets of NROs, some with occupancy near two, some with occupancy near zero, and possibly some with near single occupancy. Restricting the localization to just those orbitals with nearly double occupancy gives localized orbitals which all, of necessity, have occupations near 1.9. These orbitals will be localized into the desired chemical interpretations as lone pairs, two center bonds, and possibly multicenter bonds. Localized correlating orbitals for these electron pair orbitals can be found by localizing the nearly vacant NROs. Finally, any singly occupied orbitals can be localized. Note that these localizations restricted to part of the NRO space must be of the intrinsic type.

A second procedure, less arbitrary than dividing the natural orbitals by occupation, but more difficult to apply, starts from the atom localized orbitals. The initial step is to obtain one center pair orbitals by diagonalizing each atomic submatrix of the density matrix.

Any orbitals with occupation near two are taken as chemically adapted, i.e., lone pairs, and are removed from further consideration. Next, every possible two atom submatrix of the density matrix transformed into the above found atom-diagonal-density orbitals is taken and diagonalized. Any orbitals found with occupation near two are two center bond orbitals. These bond orbitals are symmetrically orthogonalized, and are removed from further consideration. If any orbitals with occupation significantly greater than zero remain, multicenter bonding is present. Three center bonds are found by diagonalizing all possible three atom submatrices and taking any orbitals with occupancy near two, after symmetric orthogonalization. This process is repeated until any possible bonds over even more centers are found. The remaining orbitals all have occupancies near zero. It is possible to form localized correlating orbitals from these, perhaps by energy localizing them. This could also be accomplished by forming the orthogonal complement to the above found two center bonds, etc.

These two schemes may not lead to MOs which are the most revealing. For any particular molecule, chemically adapted MOs may be found in some other fashion. In any case, the ability to extract such chemically significant orbitals from the FORS wavefunction is a very important feature of the FORS model.

F. Examples of Orbital Types

Examples of the various kinds of orbitals described in the preceding sections are illustrated for some of the full valence space type FORS wavefunctions of diatomics given in Chapter II. The molecules

chosen are NH, F_2 and CO. The orbitals are illustrated by contour plots giving the orbital amplitude on a planar cross section of the orbital. Each contour line represents 0.05 $bohr^{-3/2}$. Solid and dashed contours represent positive and negative amplitudes, respectively. Nodes or zeros are shown by widely spaced dashed curves.

Three different orbital sets are given for NH in Fig. 3.1. The core MO 1s and valence MO π_x (and the π_y not shown) span one dimensional orbital spaces and are, thus, the same for all three orbital sets. The first set of valence σ orbitals shown are the natural orbitals, which are, generally speaking, delocalized over the molecule. The electron occupancies of these orbitals have been extremized; there are two nearly doubly occupied orbitals and one nearly vacant correlating orbital.

The second orbital set for NH was localized according to the energy localization procedure of Edmiston and Ruedenberg (1963, 1965). The localization yields orbitals that are spatially isolated, so that sp hybrids are found on the nitrogen atom. These two hybrids have a substantial offdiagonal bond order of 0.3999. Diagonalization of the 2 x 2 nitrogen σ block gives two directed localized reaction orbitals, with occupancies of 1.9960 and 1.0239. This diagonalization increases the s character of the lone pair hybrid pointing away from the H atom, and increases the p character of the nitrogen bonding hybrid.

The final set of NH orbitals shown were localized by the projection technique. Note that these orbitals are quite like unhybridized atomic orbitals. One measure of their atomic nature is the overlap of these MOs with the corresponding free atom AO. These overlaps are shown in

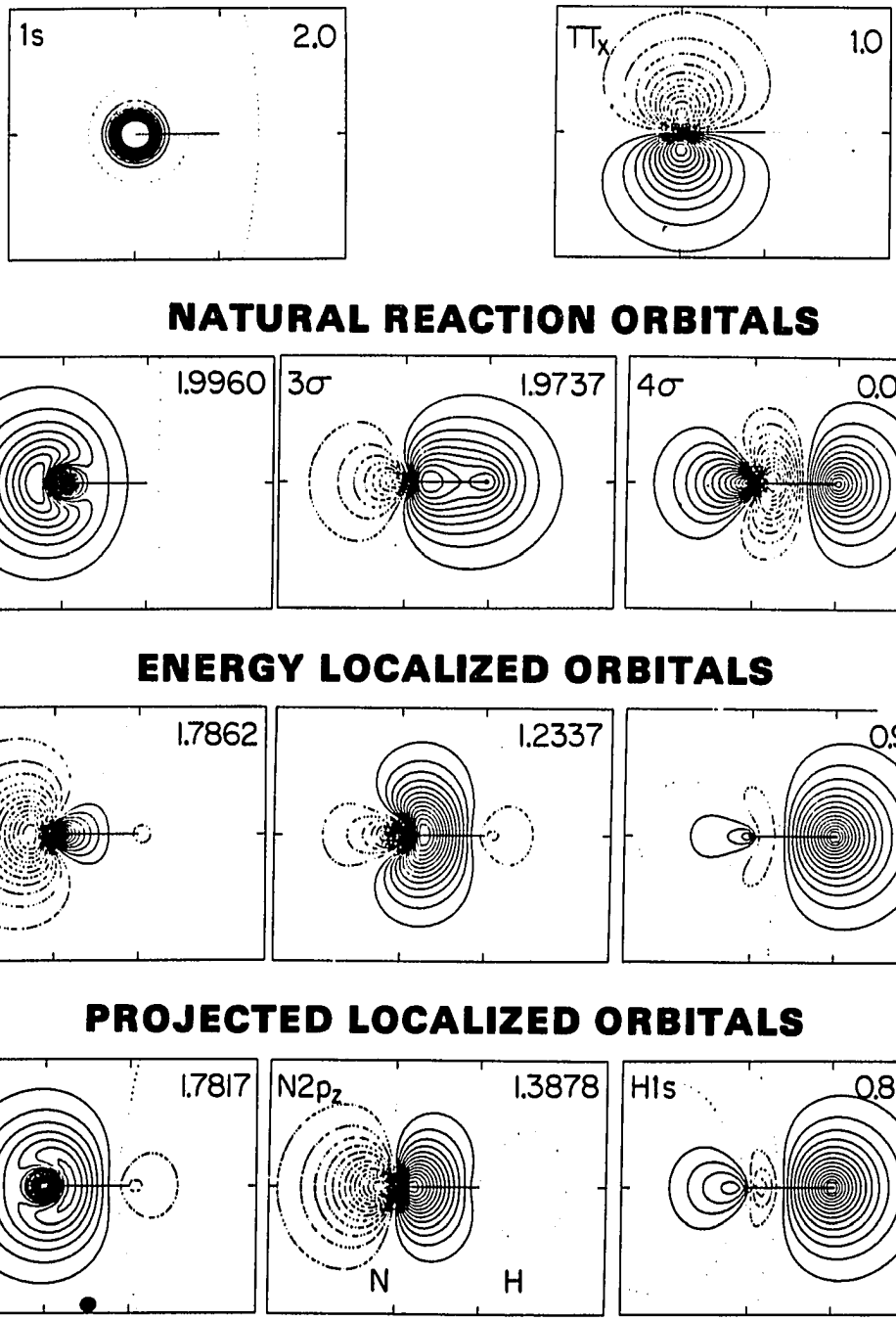


Figure 3.1. Three equivalent FORS orbital sets for the ground state of imidogen. Numbers indicate electron occupancies

Table 3.1. Inspection of the hydrogen orbitals shows that the energy and projection localization methods accomplish the localization onto atoms slightly differently.

For F_2, the natural orbitals are shown in Fig. 3.2. The projected localized orbitals are shown in Fig. 3.3. These projected orbitals are even more atomic in character than those of NH. The overlap of these MOs with the corresponding AOs is given in Table 3.1. These overlaps can be brought even closer to unity if the symmetric orthogonalization is not performed. These nonorthogonal orbitals are not shown, but their overlaps with the corresponding AOs are included in Table 3.1, and are closer to one than for the orthonormalized PLOs.

The local density submatrix on each fluorine for the orthogonalized PLOs is diagonal with respect to the π-PLOs, but not with respect to the σ-PLOs, σ(F2s) and σ(F2pz). Diagonalization of this 2 x 2 local σ density matrix yields two hybrid-type PLOs with the compositions

$$\sigma(F\ell p) = 0.980426\ \sigma(F2s) - 0.196887\ \sigma(F2pz)$$

$$\sigma(Fbo) = 0.196887\ \sigma(F2s) + 0.980426\ \sigma(F2pz)$$

on the left fluorine and the mirror images on the right fluorine. The occupation numbers of these two MOs are 1.994 and 1.006, respectively, as compared with 1.956 and 1.046 for the original 2s and 2pz type PLOs, showing that the left and right $\sigma(F\ell p)$ orbitals have lone pair character and that the left and right σ(Fbo) orbitals have bonding character. The total σ and π density matrices for these atom-adapted FORS MOs is given in Table 3.2. The only large offdiagonal element of this matrix

Table 3.1. Overlap integrals between FORS PLOs and free atom AOs

NH		F_2		CO	
N1s	0.9996	orthonormal PLOs		C1s	0.9994
N2s	0.9557	F1s	0.9994	C2s	0.9237
N2pz	0.9681	F2s	0.9900	C2pz	0.8630
N2px	0.9987	F2pz	0.9818	C2px	0.9437
H1s	0.8886	F2pπ	0.9982	O1s	0.9999
		nonorthogonal PLOs		O2s	0.9090
		F1s	0.9993	O2pz	0.9471
		F2s	0.9991	O2px	0.9722
		F2pz	0.9949		
		F2pπ	0.9997		

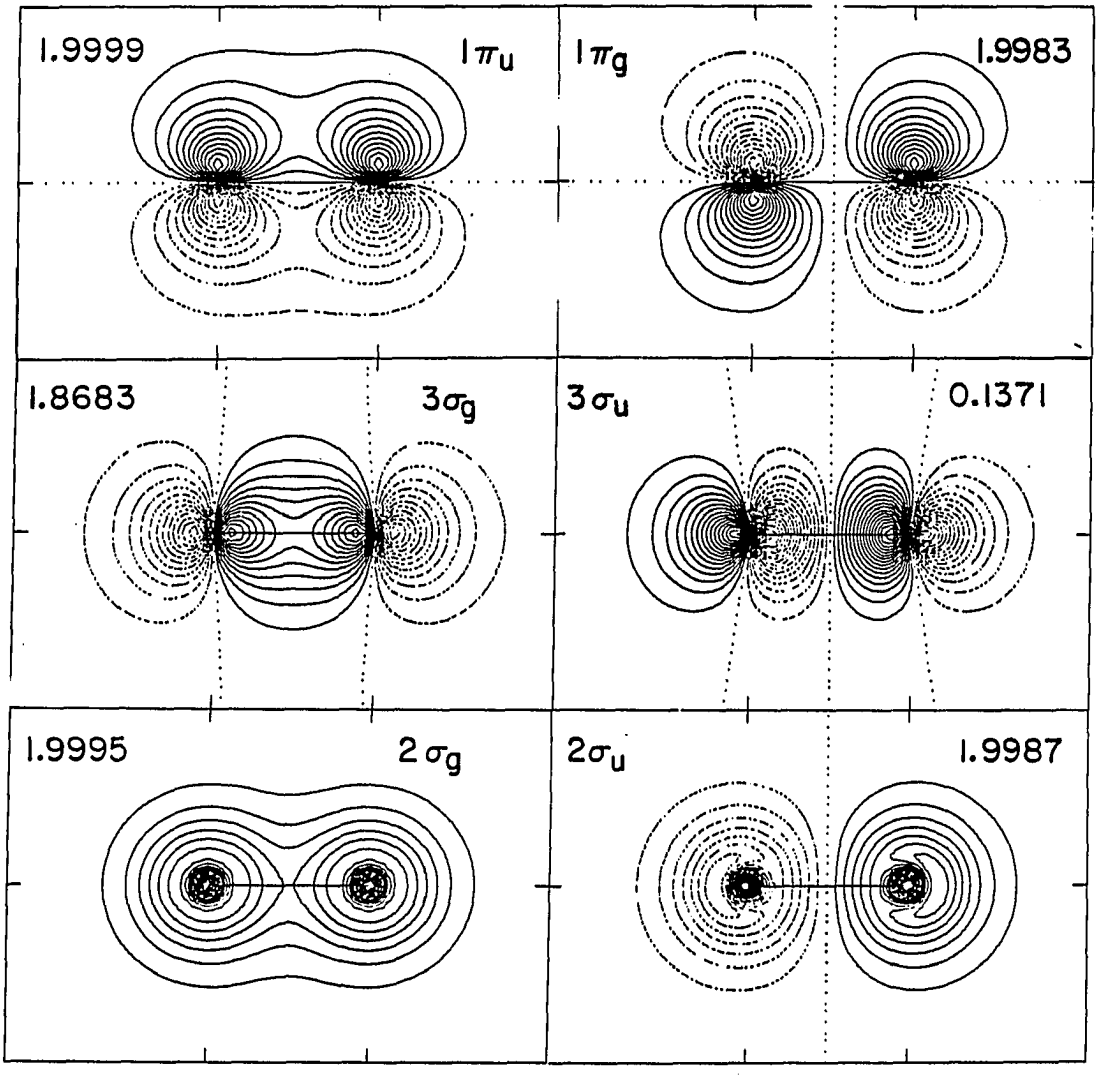


Figure 3.2. Natural reaction orbitals for F_2. Numbers give electron occupancies

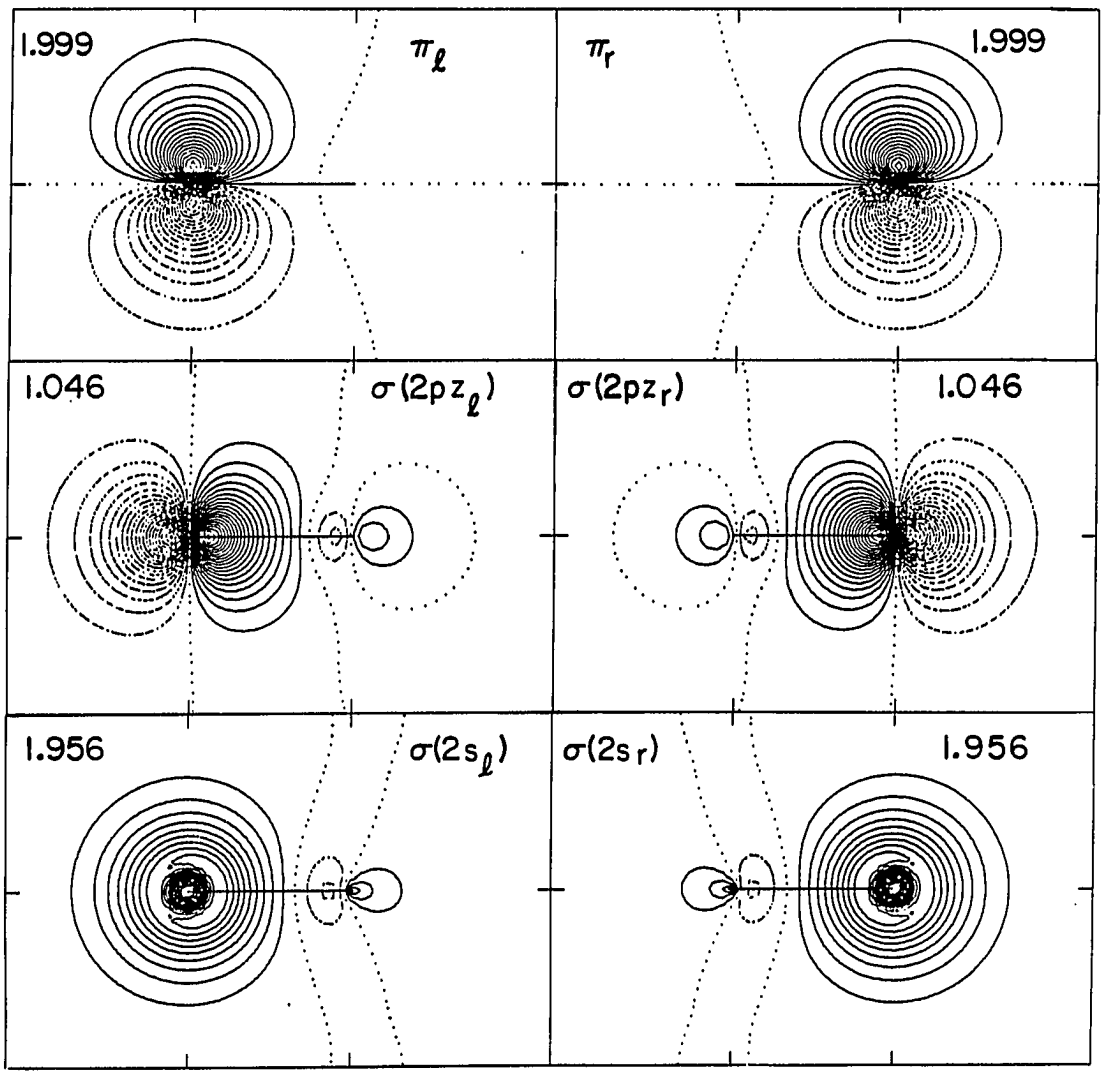


Figure 3.3. Projected localized orbitals for F_2. Numbers give electron occupancies

Table 3.2. Bond order matrix for F_2 following atomic diagonalization

		Left		Right	
		σ(Fℓp)	σ(Fbo)	σ(Fℓp)	σ(Fbo)
Left	σ(Fℓp)	1.994	0	-0.004	0.034
	σ(Fbo)	0	1.008	0.034	0.870
Right	σ(Fℓp)	-0.004	0.034	1.994	0
	σ(Fbo)	0.034	0.870	0	1.008

	π(F,left)	π(F,right)
π(F,left)	1.999	0.001
π(F,right)	0.001	1.999

indicates the only strong interaction is between the left and the right σ(Fbo) MOs.

One can now form the binding and antibinding combinations of these left and right atom-adapted bonding MOs,

$$\sigma(F_2bo) = [\sigma(\text{left Fbo}) + \sigma(\text{right Fbo})]/\sqrt{2} \quad ,$$

$$\sigma(F_2ab) = [\sigma(\text{left Fbo}) - \sigma(\text{right Fbo})]/\sqrt{2} \quad .$$

The resulting chemically adapted σ-type FORS MOs have the density matrix given in Table 3.3. This density matrix is very nearly diagonal. Complete diagonalization of the density gives the natural orbitals, which have only slightly more extremized occupation numbers, at the expense of delocalization over the diatom. Plots of the chemically adapted MOs are shown in Fig. 3.4. Eight of these MOs represent inner shells (not shown) and lone pairs with occupancies near 2. The ninth FORS MO represents the sigma bond and the last FORS MO correlates the electrons in the sigma bond. It is apparent that this choice of FORS MOs is a chemically motivated compromise between natural and completely localized orbitals. The lone pair MOs and the bonding MO clearly correspond to the electron dashes in the Lewis structure $|\underline{\overline{F}} - \underline{\overline{F}}|$. It is also apparent that the chemically adapted orbitals with occupation numbers near 2 are similar in character to localized SCF MOs. The former are, thus, a generalization of the latter. The central dash in the Lewis structure of F_2 is represented by one localized bond orbital in the SCF wavefunction, while this electron pair occupies a bond and correlating antibond in the FORS wavefunction. It is this increase in the

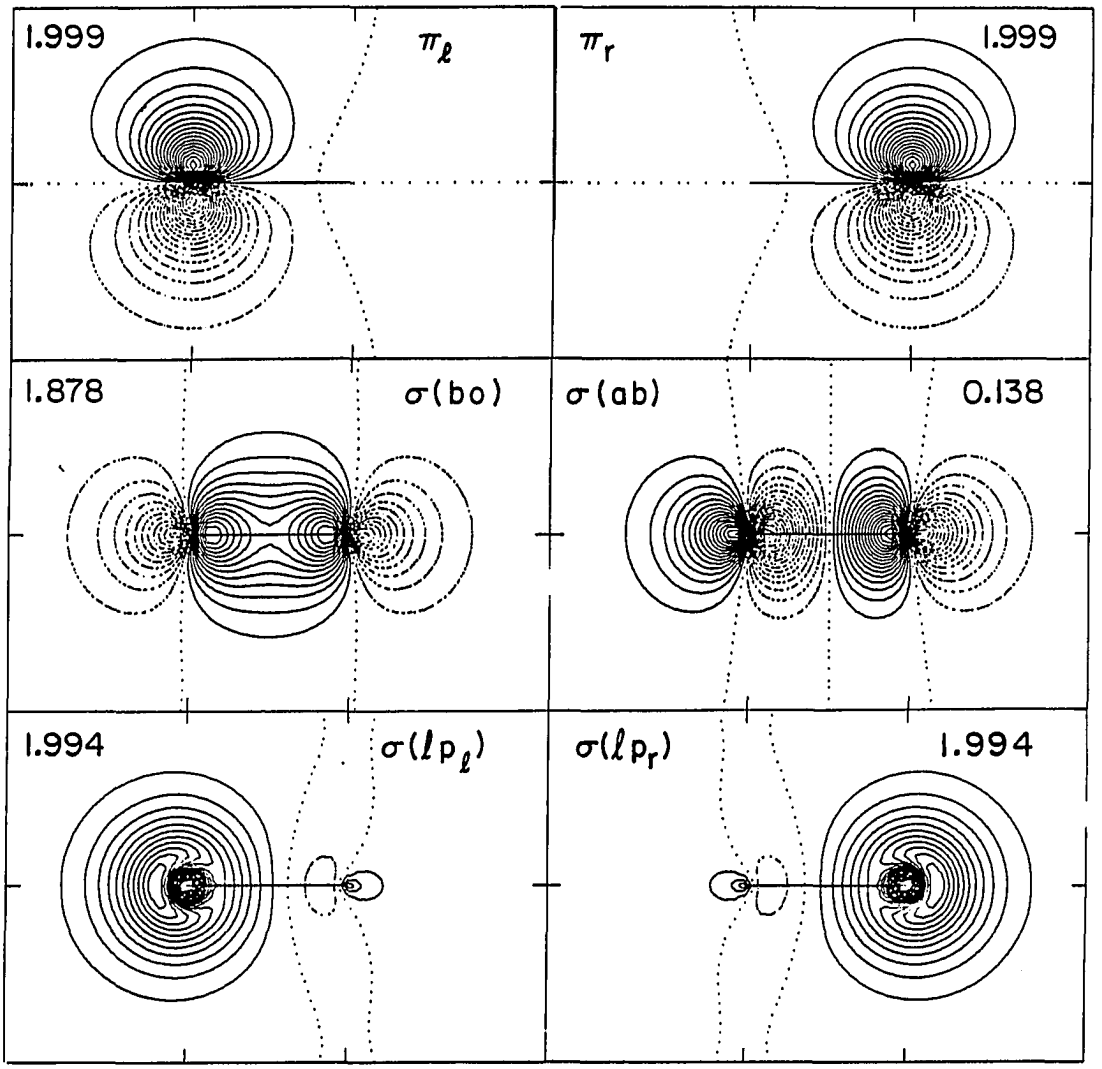


Figure 3.4. Chemically adapted orbitals for F_2. Numbers give electron occupancies. Compare nearly doubly occupied orbitals with the Lewis structure for F_2

Table 3.3. Bond order matrix for F_2 chemically adapted MOs

	σ(left Fℓp)	σ(right Fℓp)	σ(F_2bo)	σ(F_2ab)
σ(left Fℓp)	1.994	-0.004	0.024	0.024
σ(right Fℓp)	-0.004	1.994	0.024	-0.024
σ(F_2bo)	0.024	0.024	1.878	0
σ(F_2ab)	0.024	-0.024	0	0.138

sophistication with which the central dash is treated that accounts for the improvement of the FORS over the SCF wavefunction.

The final example of orbital types is CO. The natural orbitals are shown in Fig. 3.5, with the carbon atom at the left. Since CO is a diatomic, the natural orbitals are delocalized only over two atoms, and thus are to that extent chemically adapted. These natural orbitals explain very nicely the ligand chemistry of CO. The carbon end of CO bonds to metal atoms by forming σ donor and π^* acceptor bonds. The σ donation is from the nearly filled 4σ natural orbital, while the nearly vacant 2π natural orbital serves as the π^* acceptor.

The projected localized orbitals of CO are shown in Fig. 3.6. One glance shows that these orbitals, particularly on the carbon, are less atomic than the NH or F_2 PLOs. This is born out by the overlap integrals between PLOs and corresponding AOs given in Table 3.1. The populations of the two atoms are interesting. Summing the occupations of the PLOs on each atom gives 5.965 e^- on carbon and 8.035 e^- on oxygen. For comparison, a conventional Mulliken population analysis gives C: 5.924 e^-, O: 8.076 e^-. Note that the small computed dipole moment for the CO FORS function, which unlike atomic populations is a physical observable and was given in Chapter II, has the opposing sense C^-O^+.

G. The FORS Model Reexamined

The presentation of the FORS model in Chapter II and III has of necessity been in a linear fashion. The reasoning behind the model is actually more of a full circle. In Chapter II, the number of orbitals

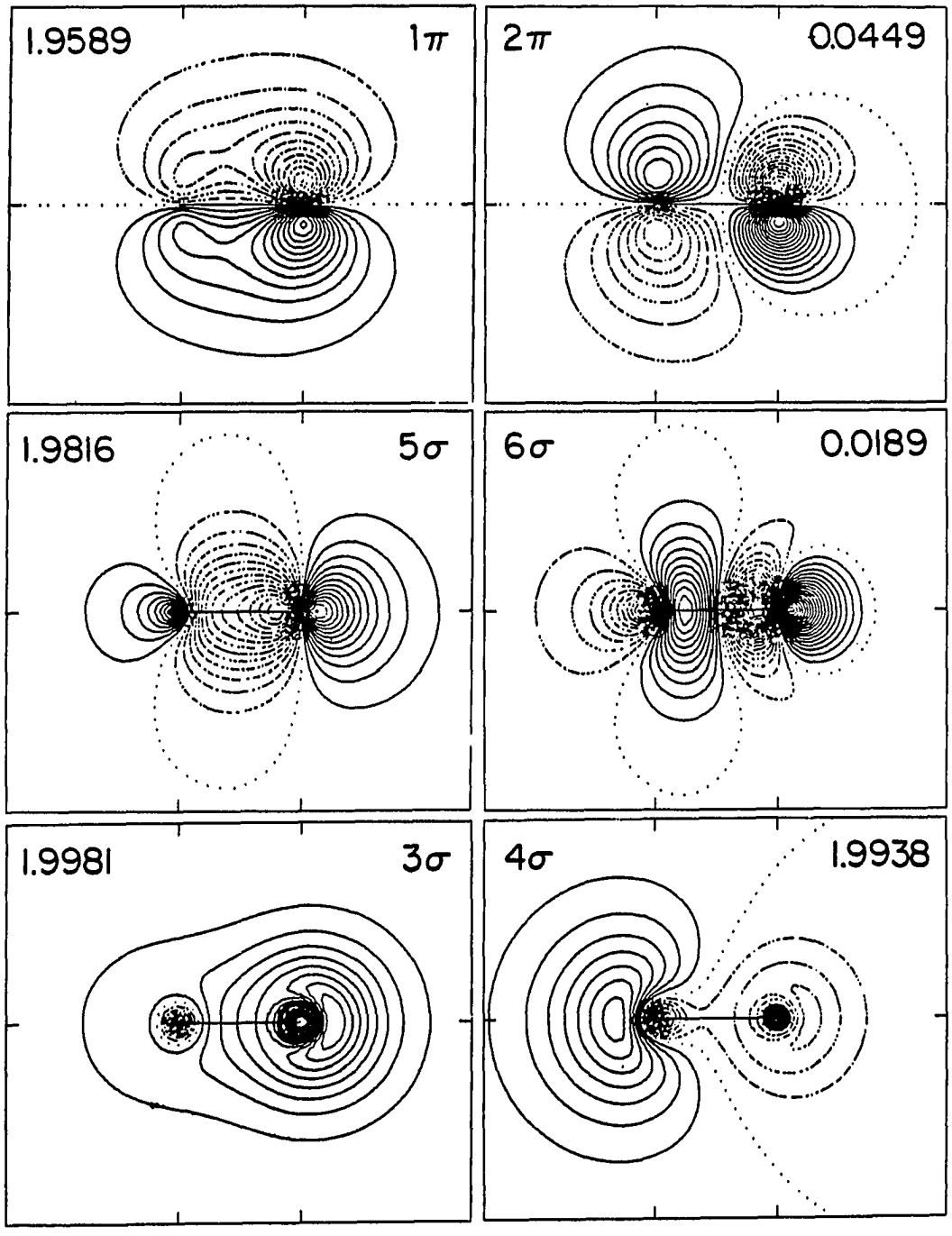


Figure 3.5. Natural reaction orbitals for CO. Numbers give electron occupancies

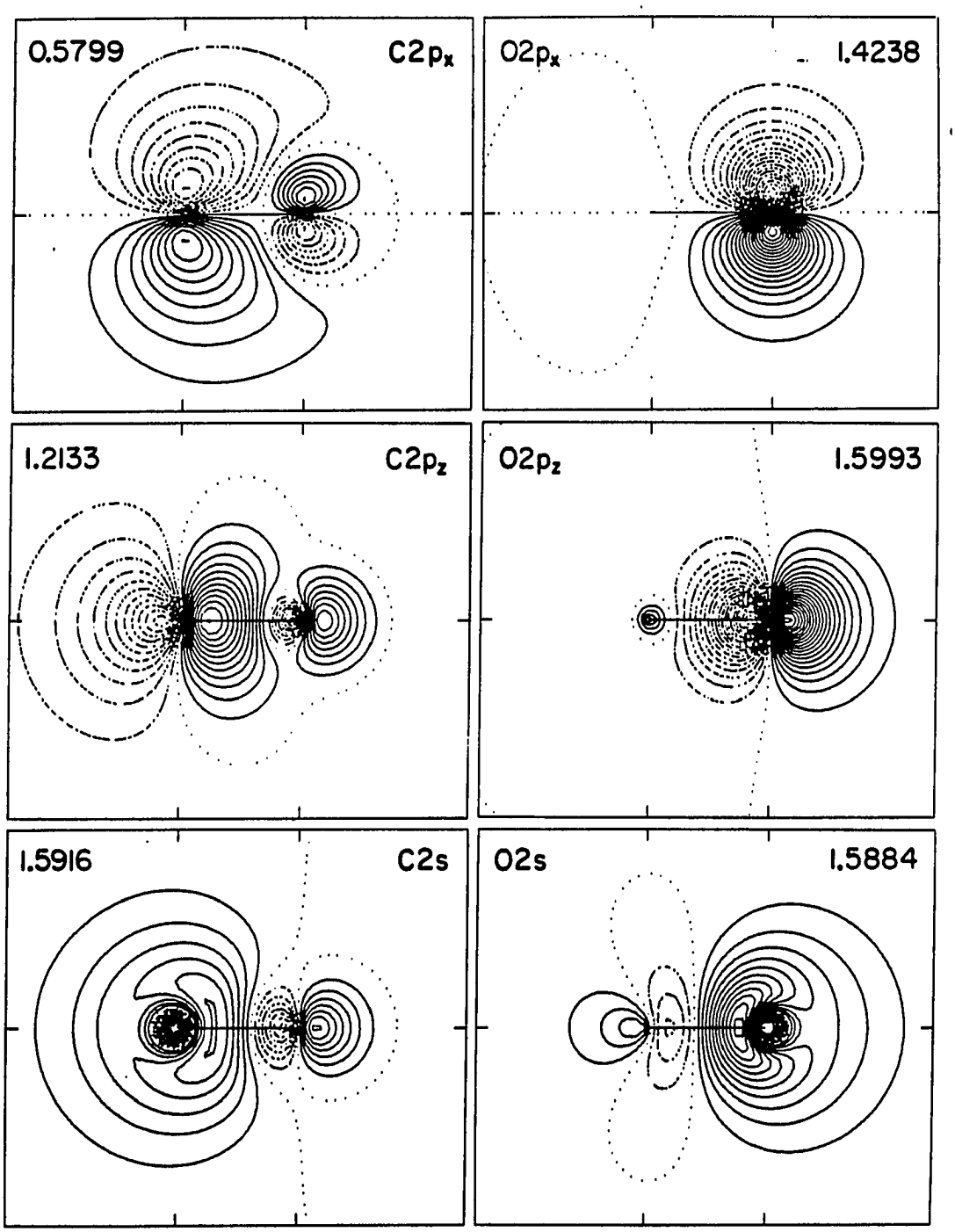


Figure 3.6. Projected localized orbitals for CO. Numbers give electron occupancies

in the full valence space type FORS function was defined as the number of minimal basis set AOs present. As a consequence, the localization schemes described in this chapter give as many localized MOs on each atom as in that atom's formal minimal basis. It is just to attain this result that the number of FORS orbitals was chosen so. In Chapter II, the atomic MBS orbitals were shown to give reasonably accurate natural orbitals, but that a significant amount of energy can be recovered by optimization in an extended basis set. In this chapter, the optimal localized MOs are shown to possess almost, but not quite, unit overlap with the free atom MBS. The atomic MBS AOs are almost optimal MOs; the optimal MOs are almost identical to the MBS.

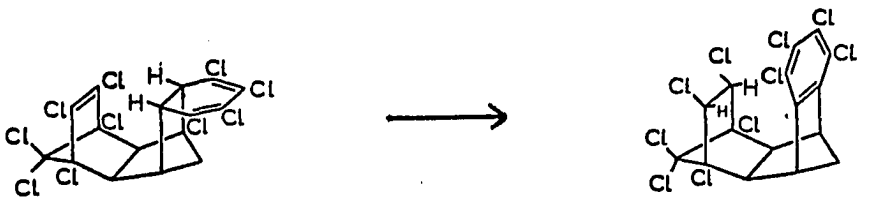
The analysis in these two chapters confirms the dominance of the atomic MBS in molecular wavefunctions. The FORS model uses the simple concept of the atomic minimal basis to define a sophisticated molecular wavefunction. As a consequence, a FORS wavefunction, with a quantitative reliability as shown in Chapter II, is open to qualitative interpretation, as demonstrated in this chapter.

IV. CONCERTED DIHYDROGEN EXCHANGE BETWEEN ETHANE AND ETHYLENE. SCF AND FORS CALCULATIONS OF THE BARRIER

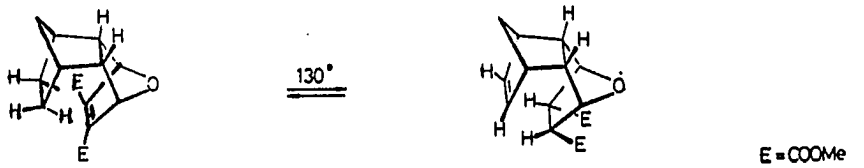
A. Introduction

The dihydrogen exchange reaction detailed in this chapter is an application of the FORS model to a reaction of greater chemical significance than the dissociation of diatomics. The investigation of this reaction has been published by Feller, Schmidt and Ruedenberg (1982). The SCF geometry optimization and vibrational analysis calculations described below were performed by David Feller.

The least-motion concerted transfer of two hydrogen atoms from eclipsed ethane to ethylene is symmetry-allowed, and has been explicitly discussed as a paradigm by Woodward and Hoffmann (1971) and by Goddard (1972). Rye and Hansen (1969) have conjectured such a process to play an intermediary role in the hydrogenation of ethylene over a metal catalyst, supposing that the adsorbed ethylene might have a structure similar to that of ethane. Doering and Rosenthal (1967) have observed dihydrogen exchange from cis-9,10-dihydronaphthalene to various olefins. Mackenzie (1965, 1969) has observed intramolecular hydrogen transfer in isodrin derivatives, i.e.:



A more recent observation of this type of transfer is

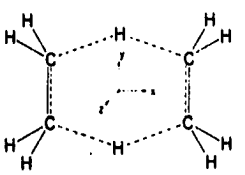


by Hagenbuch et al. (1981).

The analogous hydrogenation of olefins by diimide has long been recognized for its versatility and stereospecificity. Miller (1965) and Hunig et al. (1965) have given reviews of olefin hydrogenation by diimide. Gas phase kinetic studies by Vidyarthi et al. (1974) and Willis et al. (1977) confirm the mechanism of concerted dihydrogen transfer from diimide to olefins.

Nonetheless, the theoretical investigation reported here reveals the existence of a substantial barrier for the concerted exchange of hydrogens from ethane to ethylene.

In agreement with the reasoning by Woodward and Hoffmann (1971), a concerted movement of both hydrogens is assumed, maintaining C_{2v} symmetry throughout. A further assumption is that the activated complex of highest energy occurs for the transition state of D_{2h} symmetry:



That this geometry is indeed a saddle point on the energy surface is verified afterwards.

In the present case, the "conservation of symmetry" is equivalent to the statement that a single-determinant Hartree-Fock SCF wavefunction can be used to represent the system at all stages of the reaction in the sense that the occupied MOs of the reactants deform continuously and smoothly into the occupied MOs of the products. Accordingly, ab initio calculations at the SCF level were performed for the reactant/product geometry and for the transition state geometry. Subsequently, the validity of this approximation and, hence, the applicability of the concept of symmetry-allowedness is verified by carrying out Full Optimized Reaction Space (FORS) multiconfiguration SCF (MCSCF) calculations which allow for the dominant electron correlation effects on the calculated barrier. Orbital analysis of the SCF and FORS wavefunctions serve to deepen the understanding of the electronic rearrangements.

B. Calculational Details

Three different basis sets, taken from Schmidt and Ruedenberg (1979), were used during this investigation. The geometry optimization of the complex was performed with the smallest basis, denoted basis A, consisting of a C(6s3p/4s2p), H(3s/2s) QBO set, with the hydrogen functions scaled 1.1. Once the optimal geometry was found, larger bases were employed. Basis B, of double zeta quality contained an increased number of primitives, namely C(10s5p/3s2p), H(4s/2s) with the hydrogen unscaled. Basis C, of double zeta plus polarization quality, contains basis B and a set of d functions on carbon and p functions on the

reacting hydrogens, with exponents $\zeta_D = 0.75$ and $\zeta_p = 1.0$, taken from Dunning and Hay (1977). This final basis contains 170 PAOs and 86 QBOs.

The Hartree-Fock SCF calculations on ethane, ethylene, and the intermediate complex were performed using all three bases. The Full Optimized Reaction Space (FORS) MCSCF calculations were undertaken with basis B.

All calculations, except the normal mode analysis, were performed using the ALIS system for quantum chemical molecular calculations. The calculation of the normal modes and their frequencies was done with the GAMESS (General Atomic and Molecular Electronic Structure System) program system of Dupuis, Spangler and Wendoloski (1980), using the standard 3-21G basis of Binkley et al. (1979).

C. Results and Discussion of SCF Calculations

1. Geometry

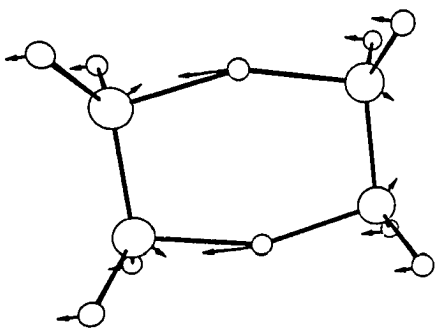
The determination of the transition state of the reaction considered here is greatly simplified by the fact that it has D_{2h} symmetry. Using basis A, all geometric parameters of the activated complex were optimized within this symmetry by SCF calculations. The optimal geometry is presented in Table 4.1. The CC bonds of the complex are intermediate in length (1.42 Å) between ethylene (1.33 Å) and ethane (1.57 Å). The four central CH bonds are quite long (1.38 Å), but the peripheral CH bond remains 1.09 Å. The peripheral hydrogens are bent back 29.0 degrees, again intermediate between ethylene (0°) and ethane (51.3°).

Table 4.1. Optimal geometry of D_{2h} ethane-ethylene complex

Atoms	x[a]	y[a]	z[a]
C	±2.555	±1.341	0
H (reacting)	0	±1.878	0
H (peripheral)	±3.095	±2.315	±1.729

[a]Coordinates of atoms in units of bohrs (1 bohr = 0.5292 Å) and referring to the coordinate system shown in the text.

That this optimized geometry of the intermediate complex is indeed a saddle point on the molecular energy hypersurface was verified by calculating all normal modes and their frequencies. Diagonalization of the mass weighted Hessian matrix yielded exactly one negative force constant which is the condition for a transition state. The corresponding imaginary frequency was found to be $2183i$ cm^{-1}, indicating a rather sharply curved energy surface in the direction of the reaction coordinate. The latter is the b_{3u} normal mode which consists of the set of atomic displacements shown here:



It is apparent that the left hand side of this figure depicts the deformation towards ethane (decrease of the HCC angle, increase of the CC distance, decrease of the distance between carbon and its neighboring central hydrogen), whereas the right hand side depicts the deformation towards ethylene (increase of the HCC angle, decrease of the CC distance, increase of the distance between carbon and its neighboring central hydrogen). The center of mass is, of course, at rest. If the central hydrogens are thought of as belonging to the right hand part of the complex, then the indicated displacements describe the entrance of the reactants into the transition state. If the central hydrogens are imagined to belong to the left hand part of the complex, then the

picture describes the exit out of the transition state towards the products. Clearly, this normal mode describes the concerted exchange of both hydrogens.

For ethane, the experimental geometry of Kuchitsu (1966) was assumed, while for eclipsed ethane, the SCF optimized geometry of Clementi and Popkie (1972) was taken.

2. SCF picture of the electronic structure

The low lying molecular orbitals, which are essentially involved in this reaction, arise from one sp^3-type hybrid on each carbon and a 1s-type orbital on each of the exchanging hydrogens. Each of these six orbitals contributes one electron to the reacting set. Because of the quasi-hexagonal arrangement of these atomic orbitals, the molecular orbitals of the activated complex are similar to those of the D_{6h} benzene π system. Of course, the hydrogen and carbon orbitals are inequivalent, lowering the symmetry to D_{2h}, and lifting the degeneracy of the benzene e orbitals, as illustrated in Fig. 4.1. Nonetheless, the nodal structures of the transition state orbitals are quite similar to those of benzene. For the SCF calculations, the three lowest of these molecular orbitals are doubly occupied, and the three highest are vacant. This quasi-aromatic character of the transition state lends support to the view that the reaction is symmetry-allowed. It should be noted, however, that the highest occupied orbital ($4b_{3u}$) in the activated complex is the π orbital of ethylene delocalized antisymmetrically between the left and right CC bonds and, hence, represents an antibonding combination of the left and right π MOs.

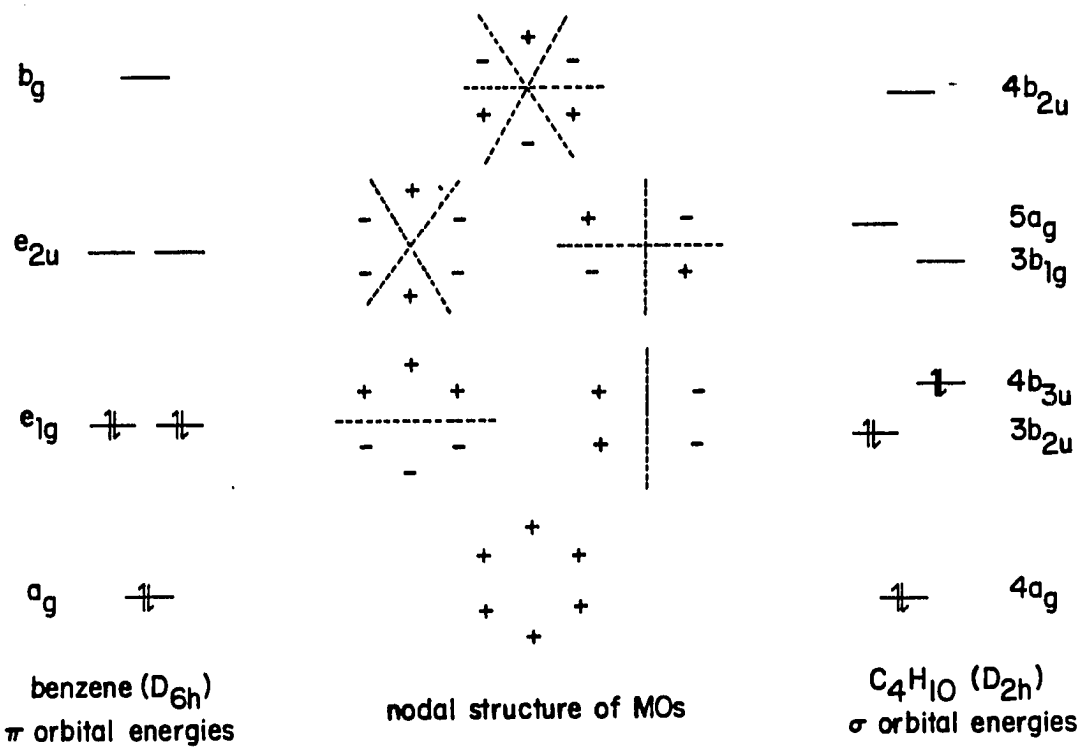


Figure 4.1. Comparison of the nodal structure of the reactive MOs of the transition state with that of the π MOs of benzene

In addition to the three occupied MOs in the reaction zone which were just discussed, there are fourteen additional occupied MOs corresponding to four carbon inner shells, two CC sigma bonds, and eight peripheral CH bonds. When canonical SCF orbitals are determined, then the ten nonreacting valence MOs get mixed with the three occupied reaction MOs discussed in the preceding paragraph. In order to identify clearly the essential changes along the reaction path, it is, therefore, necessary to generate molecular orbitals that are localized inside and outside the reaction zone. We accomplished this separation into reacting and nonreacting MOs by performing the following partial Edmiston-Ruedenberg (1963, 1965) localizations. In ethane, all valence MOs were localized yielding one CC sigma bond and six CH sigma bonds. In ethylene, the five sigma valence MOs were localized yielding one CC sigma bond and four CH sigma bonds, but the π MO was left in canonical form. For the activated complex, twelve MOs were localized, namely all valence MOs excepting only the $4b_{3u}$ MO which was left in canonical form. This localization of the transition state yielded two equivalent CC sigma bonds, eight equivalent peripheral CH sigma bonds, and two equivalent CHC three-center bonds describing the concerted hydrogen transfer. The canonical $4b_{3u}$ MO describes the shift of the π bond from one CC atom pair to the other. Contour plots of the three reacting MOs mentioned last are shown in Fig. 4.2 for the reactants/products and the transition state illustrating the progress of the reaction.

An alternative SCF MO interpretation of the transition state is obtained when <u>all</u> orbitals, <u>including</u> the $4b_{3u}$ MO, are incorporated in

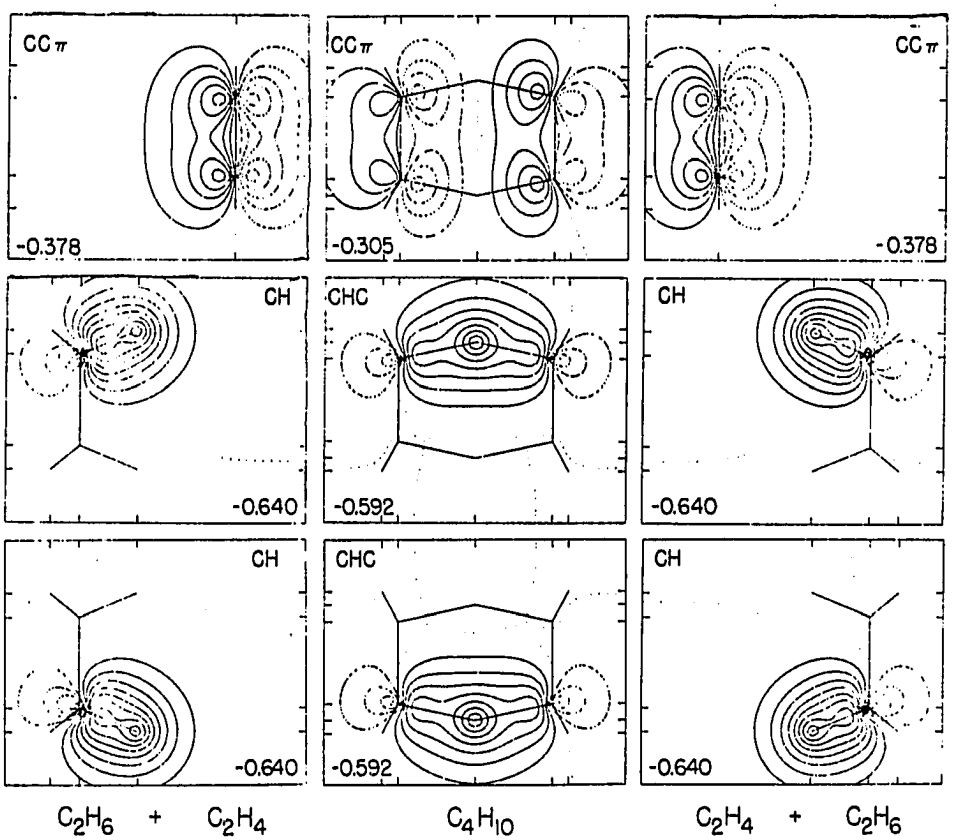
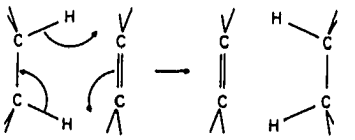


Figure 4.2. Partially localized SCF orbitals in the reaction zone. Numbers give orbital energies in hartrees

the localization process. When this is done, one obtains in the reaction zone, the MOs depicted in Fig. 4.3 rather than those shown previously in Fig. 4.2. (The SCF wavefunction formed with the MOs of Fig. 4.3 is, of course, identical to the one formed with the MOs of Fig. 4.2.) The MOs of Fig. 4.3 can be thought of as representing the intermediate stage of the continuous and concerted orbital shifts and deformations which are most easily summarized by the schematic diagram:



In the left column of Fig. 4.3, the three MOs are localized into three two-center bonds. In the middle column, each one of them is delocalized into a three-center bond covering three adjacent atoms. In the right-hand column, each MO is now relocalized into the adjacent two-center bond. This picture of an MO crawling to a neighboring site corresponds remarkably well to the idea of "electron pushing" associated by organic chemists with shorthand diagrams such as the above. Figure 4.3 demonstrates that, for the reaction at hand, this type of diagram has a rigorous quantum mechanical meaning, if it is understood to imply adiabatic orbital deformations rather than dynamic motions of electrons. It is worth noting, however, that those two intermediate transition state MOs of Fig. 4.3 which represent mixtures between a CH bond and a CC π bond (lower two plots) contain nonnegligible antibonding

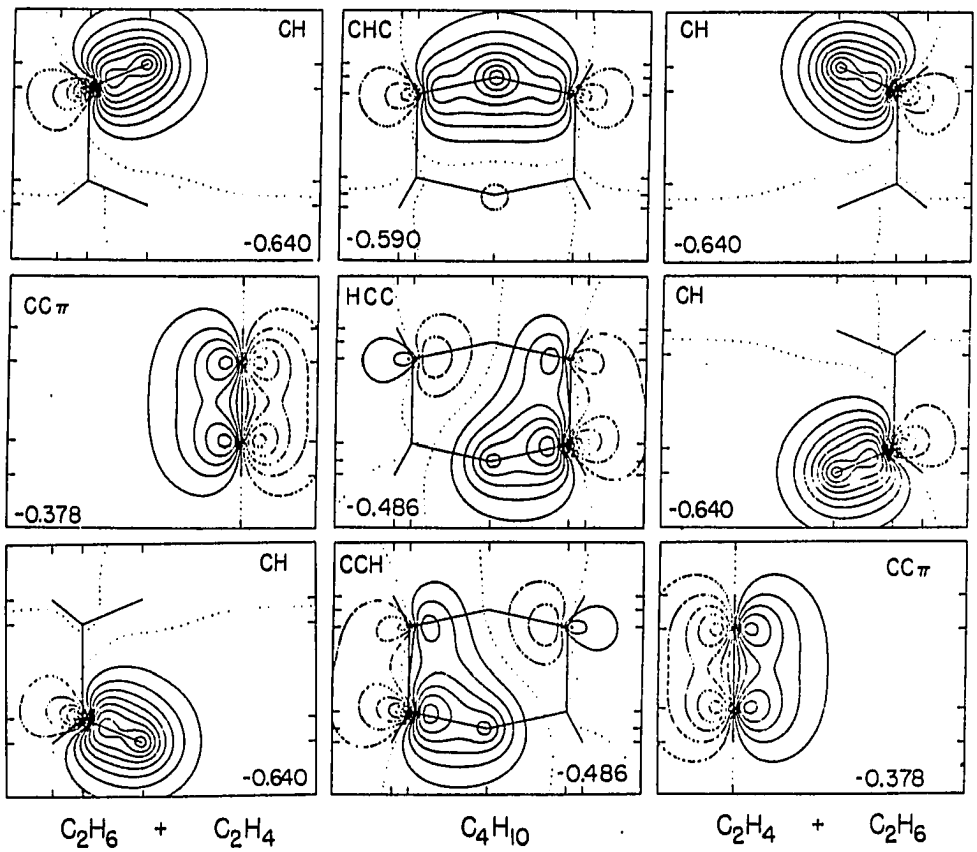
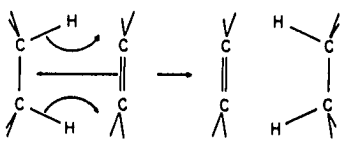


Figure 4.3. Completely localized SCF orbitals in the reaction zone. Numbers give orbital energies in hartrees

contributions at the C atom diagonally across from the main orbital lobe. These two antibonding contributions are equivalent to the previously mentioned antibonding contribution of the $4b_{3u}$ MO.

The partially localized SCF orbitals in Fig. 4.2 correspond to this "electron pushing" diagram:



3. Energies

SCF energies of ethylene, eclipsed ethane, and the intermediate complex for all three bases are reported in Table 4.2, along with the predicted reaction barriers. These barriers are underestimates because of the assumption of an eclipsed conformation for ethane as it enters the reaction coordinate. Nonetheless, they are all quite large. The computed SCF reaction barrier increases with increasing flexibility of the basis to a value of nearly 77 Kcal/mole for the largest bases considered.

That this barrier is related to the energetic changes of the three reacting MOs depicted in Fig. 4.2 can be inferred from Fig. 4.4. It exhibits quantitatively the orbital energies, i.e., the diagonal elements of the Fock matrix, of all σ-type localized valence orbitals and of the π canonical orbitals for the reactants/products and for the transition state; the connecting lines indicating the progress of the reaction. It

Table 4.2. Calculated total energies and barrier to dihydrogen exchange

Basis	Level of Calculation	C_4H_{10} (Hartree)	C_2H_4 (Hartree)	C_2H_6 (Hartree)	Barrier[a] (Kcal/mole)
A	SCF	-156.6792	-77.8036	-78.9871	70.0
B	SCF	-157.0813	-78.0078	-79.1923	74.5
C	SCF	-157.1338	-78.0333	-79.2228	76.7
B	FORS	-157.1524	-78.0377	-79.2252	69.3

[a]Barrier equals $E(C_4H_{10}) - E(C_2H_4) - E(C_2H_6)$, where 1 hartree = 627.51 Kcal/mole.

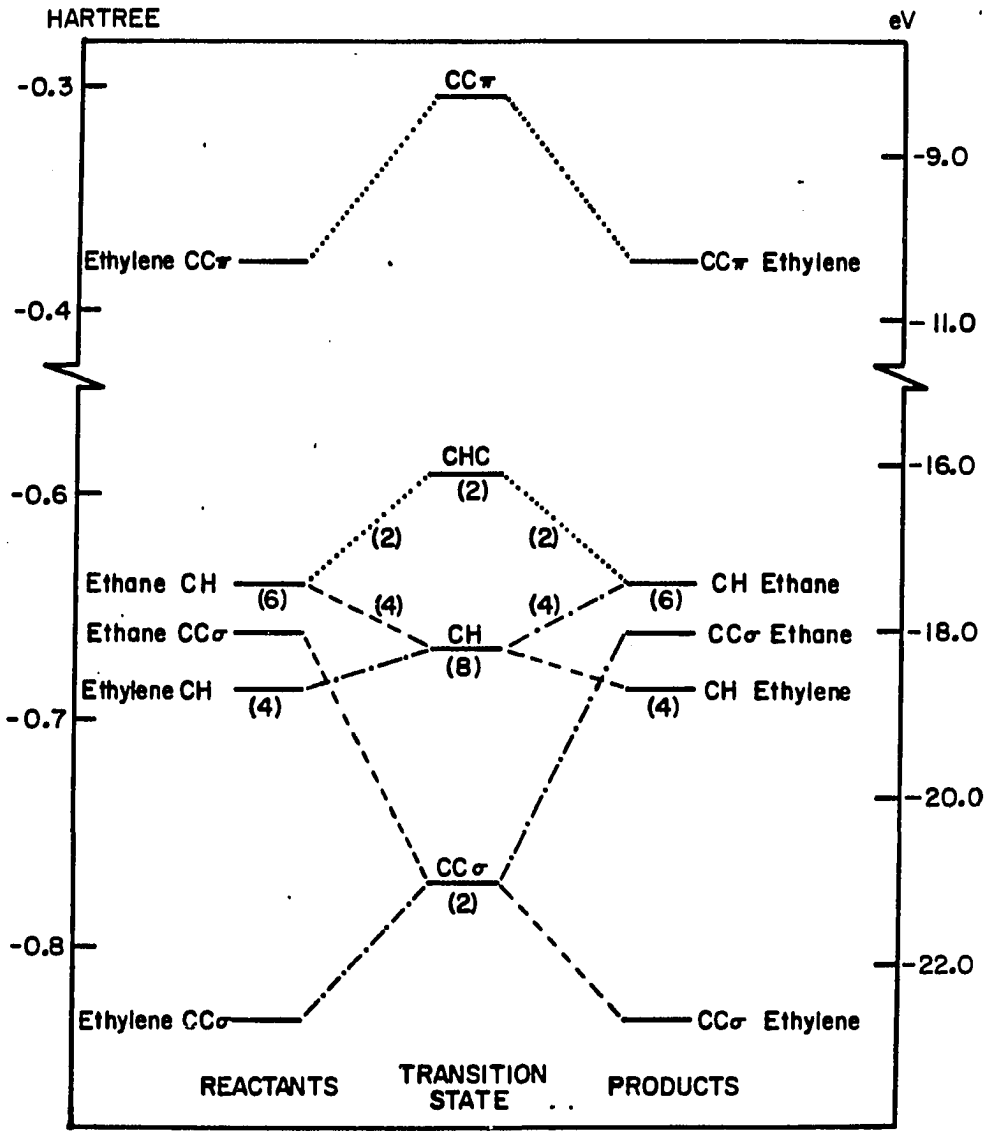


Figure 4.4. Orbital energies of SCF partially localized orbitals depicted in Fig. 4.2

is apparent that the orbital energies of the ten nonreacting MOs vary smoothly from the (higher) ethane values to the (lower) ethylene values, with the transition state having intermediate values. By contrast, the three MOs which are localized in the reaction region, i.e., the $4b_{3u}$ MO and the two CHC three-center MOs all have substantially higher values for the activated complex than for the reactants and products.

The numerical values of the MO energies of the reacting MOs are also indicated on Fig. 4.2. The orbital energies of the alternate orbitals of Fig. 4.3 are contained in that figure.

The invariant sum of <u>all</u> orbital energies is 114 Kcal/mole higher for the transition state than for the reactants/products and, thus, of the same order of general magnitude as the actual barrier. (The sum of the SCF orbital energies is well-known to differ from the total SCF energy by the electron repulsion energy.) A breakdown of this 114 Kcal/mole according to contributions from the various orbital types shown in Figs. 4.2 and 4.3 is given in Table 4.3.

D. Results and Discussion of FORS Calculations

In view of the large barrier found in the SCF approximation, two questions arise: (i) Is this result significantly changed when the wavefunction allows for correlation? and (ii) Is the reaction in fact symmetry allowed? The dominant changes due to correlation are described by the FORS model which moreover provides further insight into the electronic rearrangements.

Table 4.3. Breakdown of "sum-of-orbital-energies barrier" in terms of contributions from various orbital types

Orbital types of Fig. 4.2	ΔE^a	Orbital types of Fig. 4.3	ΔE^a
4 carbon inner shells	+10.4	4 carbon inner shells	+10.4
2 CC σ bonds	-61.8	2 CC σ bonds	-61.8
8 peripheral CH bonds	-46.9	8 peripheral CH bonds	-46.9
2 reacting CH bonds	+120.4	1 reacting CH bond	+60.2
1 reacting CC π bond	+92.1	2 bonds changing from CH to CC π or vice versa	+152.3
Total	114.2	Total	114.2

[a] $\sum_n 2\varepsilon_n(C_4H_{10}) - \sum_n 2\varepsilon_n(C_2H_4) - \sum_n 2\varepsilon_n(C_2H_6)$ expressed in Kcal/mole.

1. Wavefunctions and energies

The FORS wavefunction of the activated complex consists of the 42 A_g space orbital products, including the SCF configuration, which can be formed by distributing the six reacting electrons over the six reaction orbitals which are similar in shape to those indicated in Fig. 4.1. All possible couplings of these space orbital products with the one or two permissible spin eigenfunctions yield a total of 52 ^1A_g SAAPs which comprise the configurational basis for the FORS function. In all of them, the remaining fourteen MOs are left doubly occupied (but not frozen) as the "nonreactive core". The optimization was performed as follows. Two preliminary six SAAP MCSCF calculations followed by CI calculations for configuration selection preceded the final eight SAAP MCSCF orbital optimization. A final CI calculation was performed to obtain the final FORS energy, wavefunction, and natural reaction orbitals.

The corresponding configurational FORS bases for the reactants/products are as follows. In ethylene, it consists of the two singlet SAAPs of appropriate symmetry that can be made from the bonding π-MO and the antibonding π-MO. In ethane, it consists of the twelve singlet SAAPs of appropriate symmetry which are possible using the bonding and antibonding σ MOs in the reactive CH bonds; the other four CH bonds and the CC σ bond remaining described by doubly occupied bonding MOs. The ethylene FORS wavefunction was obtained by direct MCSCF optimization, while a preliminary three SAAP MCSCF calculation preceded the final twelve SAAP MCSCF calculation on ethane. The FORS wavefunction of the

reactant/product state is simply the normalized antisymmetrized product of the ethane and ethylene FORS wavefunctions and, hence, is a superposition of 24 SAAPs.

Only basis B was used for these calculations, and the FORS energies of all three molecules are presented as the final entry in Table 4.2. Although for each molecule the FORS energy is much lower than the SCF energy using the same basis, the difference in the electron correlation energy of the reactants/products and the intermediate complex is merely 8.3 millihartree, corresponding to a barrier lowering of only 5 Kcal/mole.

For the transition state as well as for the reactant/product state, the SCF-type configuration (in terms of natural orbitals) is dominant by far. In the 52 term transition state FORS wavefunction, this SAAP has the coefficient 0.967 corresponding to a weight of 92%. In ethane and ethylene, the SCF-type SAAPs have the coefficients 0.9903 and 0.9767, respectively, so that the coefficient of the SCF-type configuration in the 24-term FORS wavefunction of the reactant/product state is 0.9903 x 0.9767 = 0.967 as well. It is, therefore, indeed justified to consider the reaction as symmetry-allowed.

2. FORS picture of the electronic structure

As usual for any sensibly formulated FORS function, the physical interactions automatically localize the six reactive MOs in the reaction zone and the fourteen (inner shell and valence) core MOs outside the reaction zone. The core MOs can be transformed among each other by an arbitrary orthogonal transformation, and the same holds true among

the reactive MOs. Localization of the core MOs yields MOs that are very similar to the corresponding SCF MOs. In the orbital space of the six reactive orbitals, several choices of FORS MO sets are of interest.

The six natural reaction orbitals of the FORS wavefunctions for the three molecules, obtained by diagonalizing the first-order density matrix, are shown in Fig. 4.5. They furnish the orbital description which is closest to the SCF orbital description, in that is has the smallest number of dominant FORS MOs. The nodal surfaces of the natural reaction orbitals of the transition state are very similar in both number and location to those of the schematic MOs which were conjectured in Fig. 4.1. Each orbital deforms in a smooth and continuous fashion from reactants to activated complex to products, in accordance with the symmetry-allowed nature of the reaction. The symmetry-allowed nature is furthermore confirmed by the near constancy of the electron occupations of the nearly doubly occupied orbitals in Fig. 4.5, and by the smallness of the occupation numbers of the correlating orbitals in Fig. 4.5.

Localization of the six natural reaction orbitals according to the Edmiston-Ruedenberg (1963, 1965) energy criterion yields the six AO-like molecular orbitals shown in Fig. 4.6. These localized reaction orbitals have the somewhat deformed shapes of the hydrogen 1s AOs and the carbon hybrid AOs which were used conceptually to generate the symmetry-adapted full reaction space orbital set. The localized reaction orbitals represent "molecule-adapted minimal basis atomic

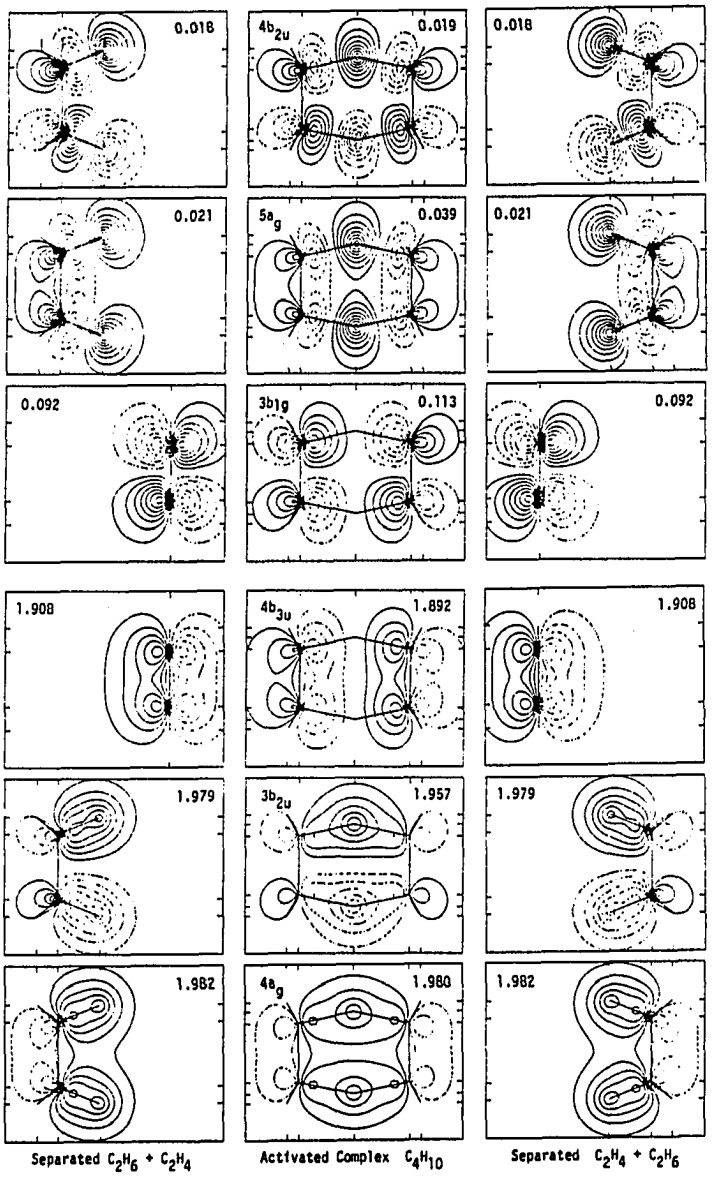


Figure 4.5. Natural reaction orbitals of the FORS wavefunctions. Numbers give electron occupancies

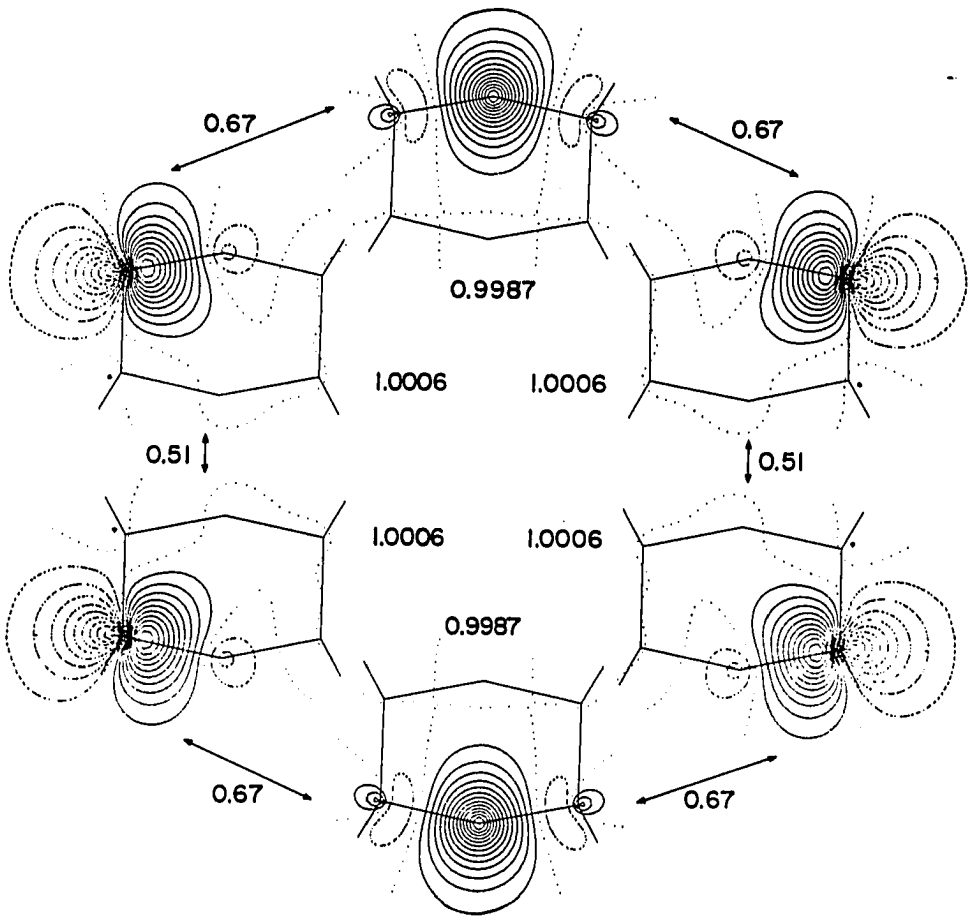


Figure 4.6. Localized reaction orbitals of FORS wavefunction of the transition state. Numbers in center give electron occupancies; numbers between orbitals are bond orders

orbitals". That is, the minimal basis set AOs can be imagined as having been deformed in such a fashion that they are directly usable as orthogonal FORS MOs. Conceptually, all other FORS MOs, such as, e.g., the natural reaction orbitals, can then be conveniently thought of as being obtained from the localized reaction orbitals by orthogonal transformations. In this manner, they can be easily visualized.

The localized reaction MOs of the FORS wavefunctions for ethane and ethylene have been very similar shapes, and Table 4.4 lists all populations and bond orders for the localized reaction orbitals of reactants, products and the transition state. For the transition state, the populations and bond orders between adjacent pairs are also indicated in Fig. 4.6. The values of the bond orders are readily understood as the sum of the contributions made by the three dominant natural reaction orbitals shown in Fig. 4.5., i.e., $p_{ij} \approx 2\Sigma_n T_{in}T_{jn}$, where $n = 4a_g$, $3b_{2u}$, $4b_{3u}$ and $\underset{\sim}{T}$ is the transformation from natural to localized reaction orbitals. (For example, the substantial negative CC bond order diagonally across the ring results from the positive $4a_g$ contribution and the negative contributions from the $3b_{2u}$ and $4b_{3u}$ MOs.) From Table 4.4, the following inferences can be drawn: (i) Each localized reaction orbital remains occupied by almost exactly one electron throughout the reaction; (ii) In ethane the two strong CH bonds interfere little with each other; (iii) The corresponding CH bonds in the transition state are weaker than those in ethane; (iv) The π-bond between the neighboring carbons, too, is weaker in the transition state than in ethylene; (v) In the transition state, all CH bonds and

Table 4.4. Populations and bond orders of localized reaction orbitals

Position	Ethane	Transition State	Ethylene
	Populations		
C atom	1.0053	1.0006	1
H atom	0.9947	0.9987	0
	Bond Orders		
CH (neighbors)	0.9787	0.6798	0
CC (neighbors)	0.0610	0.5122	0.9081
CH (next neighbors)	-0.0003	0.0006	0
CC (next neighbors)	0	-0.0018	0
CC (third neighbors)	0	-0.3778	0
HH (third neighbors)	-0.0588	-0.1128	0

CC bonds interact with each other; and (vi) There is some antibonding character across the ring which presumably contributes to the barrier.

It should be noted that the occupation numbers and bond orders given in Fig. 4.6 and in Table 4.4 are the elements of the first order density matrix in terms of the orthogonal MOs given in Fig. 4.6 and, hence, are entirely unambiguous. In view of the strong localization of these orbitals, these occupation numbers imply that, of the six reacting electrons, a charge of one resides near each of the six reacting atoms in the transition state. (These populations are not Mulliken populations, which refer to nonorthogonal and less localized atomic basis orbitals. In the present case, a Mulliken population analysis distributes the six reacting electrons as follows: reacting hydrogen = 0.930; carbon = 1.020; and peripheral hydrogen = 0.007. In view of the well-known questions regarding the definition of Mulliken populations, the meaning of these values is less definite.)

Further sets of interesting FORS MOs are chemically adapted reaction orbitals, which are defined as MOs that have nearly maximal or minimal occupation numbers and, in addition, are also localized in as small a part of the molecule as possible. For the transition state, a chemically adapted reaction orbital set is given in Table 4.5 in terms of the localized reaction orbitals. It is readily seen that three of these MOs have occupation numbers close to 2 and are very similar in shape to the partially localized SCF MOs, which were depicted in Fig. 4.2. Each of the remaining three MOs is seen to be an orbital which essentially provides "interatomic" electron correlation to one of

Table 4.5. Chemically adapted reaction orbitals corresponding to SCF MOs of Fig. 4.2 in terms of localized reaction orbitals[a]

	$\|C_1\rangle$	$\|H_1\rangle$	$\|C_2\rangle$	$\|C_3\rangle$	$\|H_2\rangle$	$\|C_4\rangle$	Occupation Numbers	
$\|CHC\rangle_1$	1/2	$1/2\sqrt{2}$	1/2	0	0	0	1.961	2
$\|CHC^*\rangle_1$	1/2	$-1/2\sqrt{2}$	1/2	0	0	0	0.039	
$\|CHC\rangle_2$	0	0	0	1/2	$1/2\sqrt{2}$	1/2	1.961	2
$\|CHC^*\rangle_2$	0	0	0	1/2	$-1/2\sqrt{2}$	1/2	0.039	
$\|CC\pi\rangle$	1/2	0	-1/2	-1/2	0	1/2	1.890	2
$\|CC\pi^*\rangle$	1/2	0	-1/2	1/2	0	-1/2	0.110	

[a]Column headings: Localized reaction orbitals, as shown in Fig. 4.6, corresponding to the indicated atoms (progressing counterclockwise, starting at one-o'clock position). Row headings: Chemically adapted reaction orbitals (labels have meanings similar to those used in Fig. 4.2).

the three dominant MOs. Another set of chemically adapted reaction orbitals is given by the orthogonal transformation in Table 4.6. It is readily verified that those MOs with occupation numbers close to 2 are very similar to the completely localized SCF MOs of Fig. 4.3, and that each of those MOs with near-zero occupation numbers again provides interatomic correlation to one of the dominant MOs. From the expansions in this table, the antibonding character across the ring mentioned earlier for the MOs |CCH> and |HCC> is particularly evident.

E. Discussion of Barrier

The calculated zero point vibrational energies of ethylene, eclipsed ethane, and the activated complex are 0.0547, 0.0787, and 0.1318 Hartree, respectively. Thus, the inclusion of zero point vibrational energies lowers the barrier by 0.0016 Hartree, which is just 1.0 Kcal/mole.

On the basis of these calculations, the reaction barrier is estimated to be 71 Kcal/mole. This estimate is obtained by subtracting the 5 Kcal/mole lowering due to the inclusion of electron correlation, using basis B, and the 1 Kcal/mole lowering due to the inclusion of the zero point vibrational energy, from the SCF barrier of 77 Kcal/mole obtained using basis C. Further calculational improvements are unlikely to lower the barrier significantly, as the largest basis used is quite flexible, and the FORS method can be expected to incorporate the pertinent changes in electron correlation which occur during the course of a reaction. Taking into account the rotational barrier in

Table 4.6. Chemically adapted reaction orbitals corresponding to SCF MOs of Fig. 4.3 in terms of localized reaction orbitals

	$\|C_1\rangle$	$\|H_1\rangle$	$\|C_2\rangle$	$\|C_3\rangle$	$\|H_2\rangle$	$\|C_4\rangle$	Occupation Numbers	
$\|CHC\rangle_1$	1/2	$1/2\sqrt{2}$	1/2	0	0	0	1.961	2
$\|CHC^*\rangle_1$	1/2	$-1/2\sqrt{2}$	1/2	0	0	0	0.039	
$\|HCC\rangle$	$1/4\sqrt{2}$	0	$-1/4\sqrt{2}$	0	1/2	$1/2\sqrt{2}$	1.925	2
$\|HCC^*\rangle$	$-1/4\sqrt{2}$	0	$1/4\sqrt{2}$	0	-1/2	$1/2\sqrt{2}$	0.075	
$\|CCH\rangle$	$-1/4\sqrt{2}$	0	$1/4\sqrt{2}$	$1/2\sqrt{2}$	1/2	0	1.925	2
$\|CCH^*\rangle$	$1/4\sqrt{2}$	0	$-1/4\sqrt{2}$	$1/2\sqrt{2}$	-1/2	0	0.075	

ethane, we finally infer a dihydrogen exchange barrier of about 74 Kcal/mole between staggered ethane and ethylene.

The origin of this large barrier is apparent from Figs. 4.4 and 4.5. While the symmetry-allowedness insures that the occupancies of the three frontier MOs remain close to 2 along the entire reaction path, it does not prevent the energies of these orbitals from changing drastically. In fact, all three have significantly higher orbital energies for the reaction intermediate than for the reactants/products. The energy increase contributed by the CHC bonds is associated with the fact that the two-electron two-center CH bonds of ethane become CHC two-electron three-center bonds; a circumstance which dilutes and weakens the binding effect and, thereby, leads to a considerable elongation of the CH bonds with a concomitant rise in energy. For very similar reasons, the spread of the two-electron two-center π-bond of ethylene over both CC bonds dilutes and weakens the bonding effect of the π-electrons, as evidenced by the lengthening of the CC bonds compared to ethylene. In addition, the CC π-bond is further destabilized by the left-right antisymmetry of the $4b_{3u}$ MO which creates an antibinding effect that increases with decreasing distance between the reactants.

For the purpose of comparing the present results with the schematic Figure 40 of Woodward and Hoffmann (1971), one can form the normalized sum and difference of the two localized CHC three-center SCF MOs of the transition state and the same linear combinations of the two corresponding CH localized SCF MOs in ethane. (Plots of the localized MOs were shown in Fig. 4.2.) The resulting SCF MOs have shapes which are

very similar to the natural reaction orbitals $4a_g$ and $3b_{2u}$, and the corresponding ethane orbitals shown in Fig. 4.5. The energies of these SCF MOs are given in Table 4.7 for the reactants/products and the transition state. In this formulation, the barrier contributions from the CHC three-center bonds are seen to be entirely concentrated in the orbital which is antisymmetric with respect to the CC bond bisectrix, whereas the symmetric orbital has nearly the same energy in the transition state as in ethane. This is so because the offdiagonal Fock matrix element between the two localized CHC MOs of the transition state is much larger than the corresponding matrix element between the CH bonds in ethane. This increase in the interference energy between the two localized CHC MOs is probably due to the shortening of the CC distances and of the distance between the central hydrogens. The orbitals $4a_g$, $3b_{2u}$, and $4b_{3u}$ correspond conceptually to the orbitals denoted by S(CH), A(CH), and S(π) in Figure 40 of Woodward and Hoffmann (1971).

Since the D_{2h} transition state has such a high energy, one might finally raise the question whether there exists a nonplanar transition state, similar to the chair or boat form of cyclohexane, whose saddle point is lower than the one considered here. While we have not made an exhaustive study of all possible geometries, we have investigated the variation of the energy for the following deformations of the planar arrangement. The four carbon atoms were kept in a plane; all distances between neighbor atoms were kept at their transition state values and all bond angles at the carbon atoms were held fixed as well. These constraints restrict the motions of the hydrogen atoms to rigid

Table 4.7. Orbital energies of SCF MOs corresponding to the natural FORS MOs of Fig. 4.5

Orbital	Ethane/Ethylene	Activated Complex	Ethylene/Ethane
$S(\pi)$, $4b_{3u}$	-0.378	-0.305	-0.378
A(CH), $3b_{2u}$	-0.604	-0.506	-0.604
S(CH), $4a_g$	-0.677	-0.679	-0.677

cylindrical rotations around the left and right CC bonds, respectively, so that the bridge hydrogens are limited to one-dimensional out-of-plane motions and the distance between the left and right CC bonds decreases somewhat. The maximum displacements considered were approximately those where the angle between the CCCC plane and the bonds to the bridge hydrogens were equal to the corresponding angles in cyclohexane, at which point the distance between the left and right CC bonds is shortened by about 0.15 Angstrom only. The boat form as well as chair form was investigated. Specific geometric parameters and the corresponding changes in the SCF energy are listed in Table 4.8. In view of the strong increase of the energy relative to the D_{2h} transition state, it seems unlikely that there exists another saddle point describing a concerted hydrogen exchange. The reason for this energetic behavior must be seen in the difference between the electronic structure of this activated complex and that of cyclohexane. Whereas the ring in cyclohexane is formed by 12 electrons using 6 bonding MOs formed from 12 molecule adapted minimal basis set AOs, the ring in the transition state at hand is formed by 10 electrons using 5 bonding MOs formed from 10 molecule adapted minimal basis set orbitals. There exists, thus, an essential delocalization of at least two electrons over all four carbons. It seems, therefore, unlikely that the CHC bridge bonds can be significantly shortened and that the antibonding effect embodied in the $4b_{3u}$ MO will disappear when the ring is made nonplanar. It rather seems likely that matters will get worse, because the π bonding character embodied <u>within</u> the left and right halves of the $4b_{3u}$ MO will be lost

Table 4.8. Energies of activated complex for out-of-plane deformations of transition state geometry

Geometric Data				Energies	
Z_B[a]	θ_B[b]	C_1C_2[c]	H_pH_p[d]	$\Delta E(Chair)$[e]	$\Delta E(Boat)$[e]
±0	0	2.704	3.276	0	0
±0.1	10.55	2.702	3.201	0.33	0.16
±0.3	29.19	2.685	3.038	2.98	1.46
±0.6	48.17	2.628	2.754	12.68	6.08
±0.9	59.18[f]	2.531	2.421	31.88	15.20

[a]Components of bridge hydrogen positions perpendicular to CCCC plane in bohr.

[b]Dihedral angle between CCCC plane and CH_BC planes in degrees (H_B = bridge hydrogens).

[c]Shortest CC distance across the ring, between left and right C_2H_4 fragments, in Angstrom.

[d]Shortest HH distance between left and right peripheral hydrogens, in Angstrom.

[e]ΔE = (SCF energy of indicated geometry) - (SCF energy of D_{2h} transition state) in Kcal/mole, calculated with 3-21G basis of Binkley et al. (1979).

[f]Note: For cyclohexane, one has $\theta_B = 54.74°$.

when both halves of the transition state assume the staggered conformation appropriate to the chair or boat form.

F. Reconciliation of High Barrier and Experimental Results

The results reported here for the prototype concerted dihydrogen exchange between ethane and ethylene imply that the intrinsic barrier (i.e., the barrier in the absence of other factors, such as discussed below) for the dihyrodgen exchange between alkyl and alkenyl carbons is about 71(74) Kcal/mole. This is surprisingly large and shows that neither the symmetry-allowed nature of the reaction nor the aromatic character of the transition state in themselves are sufficient to guarantee a low activation energy. The implication is that additional factors which stabilize the transition state are required to reduce the activation energy for dihydrogen exchange to a surmountable magnitude.

An example where such is the case is the analogous reaction between diimide and olefins, where the formation of N_2 leads to a very large lowering in the energy of the products which, in turn, is expected to depress the activation energy substantially. Thus, we have, for instance,

$$N_2H_2 + C_2H_4 \rightarrow N_2 + C_2H_6 + \text{about 90 Kcal/mole}$$

where N_2H_2 is assumed to be in the cis form. Pasto and Chipman (1979) have calculated the barrier for this reaction to be about 27 Kcal/mole. Their result is a confirmation of the results obtained in the present investigation in the following sense: Even with the assistance of a

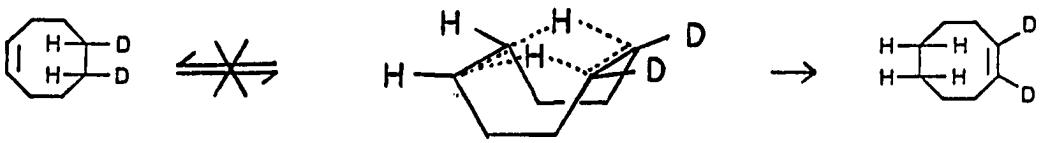
very large heat of formation of the products, there still survives a barrier as high as 27 Kcal/mole.

A carbon to carbon example is the dihydrogen transfer from cis-9, 10-dihydronaphthalene to olefins observed by Doering and Rosenthal (1967). With respect to this reaction, our results lead us to agree with the comment by Doering and Rosenthal viz. "as the driving force in the hydrogenation by diimide is very probably associated with the contribution of the heat of formation of nitrogen to the transition state, so the driving force in the transfer of hydrogen from cis-9,10-dihydronaphthalene may be dependent on the unique possibility of contributing a part of the full resonance energy of naphthalene (61 Kcal/mole) to the transition state." The intramolecular transfer noted by MacKenzie (1965, 1969) is another example where the formation of an aromatic ring lowers the activation energy.

An interesting example of an experimentally observed transfer, where the product is not much more stable than the reactant, is the intramolecular transfer observed by Hagenbuch et al. (1981). They find an equilibrium constant of around unity, so that the system is thermally neutral. However, the "cage" structure forces the ethane and ethylene like portions of the molecule to be quite close. As noted by Hagenbuch et al., this compression would provide part of the energy needed to surmount the barrier. Using the geometric assumptions of Hagenbuch et al. leads to a distance of 2.3 Å across the ring, which is much shorter than the 2.7 Å distance found for the optimal transition state between ethane and ethylene. SCF calculations with the 3-21G basis show

that bringing ethane and ethylene to 2.7 Å, with unrelaxed internal geometries, raises the energy 33 Kcal/mole. Further compression to 2.3 Å requires an additional 70 Kcal/mole. Of course, some relaxation in the ethane and erthylene molecules will reduce this requirement. Obviously, the 0.4 Å extra compression forced on the ethane and ethylene moieties in the transfer observed by Hagenbuch et al. provides an energy input that is a significant portion of the intrinsic barrier height reported. Hagenbuch et al. estimate the enthalpy of activation to be 35-59 Kcal/mole. The reduced barrier allowing the transfer to occur is due to compression of the reactants, rather than great stability of the products.

Further evidence supporting a high barrier is the failure to observe dihydrogen transfer in deuterium labeled cyclooctane (Bellamy, 1972):



No reaction is observed after 7 days at 250°C. Here the intermediate is in the "boat" conformation. As shown above, this is even less stable than the planar transition state.

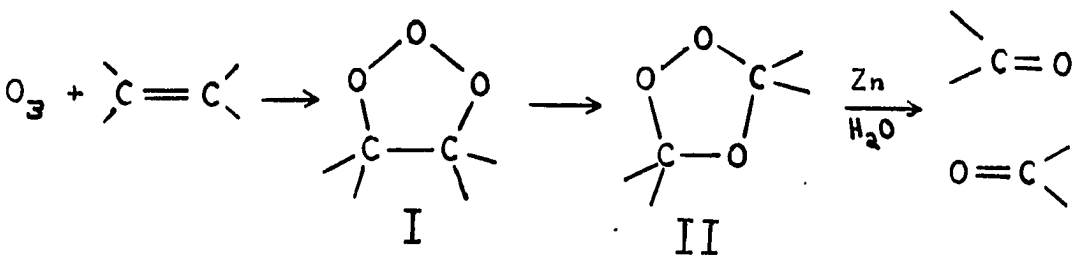
The intrinsic barrier to the Woodward-Hoffmann allowed concerted dihydrogen transfer is 71 Kcal/mole. Therefore, this type of transfer can occur only in situations where the transition state is stabilized by some factor in addition to the symmetry allowedness.

V. C_{2v} REACTIONS IN THE DIOXIRANE/DIOXYMETHANE SYSTEM: APPLICATION TO OZONOLYSIS

A. Introduction

1. Chemical background

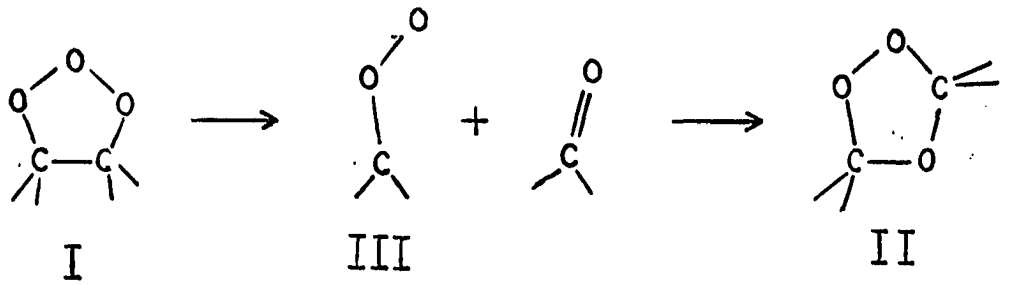
Ozonolysis of olefinic double bonds has been investigated from the turn of the century (Harries, 1903). Ozonolysis of alkene solutions is a common technique; the degradation to smaller aldehydes and ketones permits identification of the original alkene. The mechanism for ozonolysis was first proposed by Criegee (1957) and modified to account for stereochemistry by Bauld et al. (1968). The broad outline of the Criegee, or Bauld-Bailey mechanism for solution phase ozonolysis is as follows:



I II

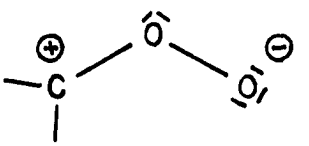
The thermochemistry has been estimated by Wadt and Goddard (1975). Structure I, known variously as the primary ozonide, molozonide, or 1,2,3-trioxolane, is about 50 Kcal/mole below the reactants. A primary ozonide has never been isolated, and its structure is still somewhat uncertain. Hiberty (1976) has found via SCF calculations that five member rings are more stable than other possibilities for I, and that an oxygen envelope is slightly more stable that a half chair conformation.

Structure II, known as the secondary ozonide, or 1,2,4-trioxolane, is readily detected in solution, and lies about 35 Kcal/mole below I. The Criegee mechanism for the conversion of I to II is

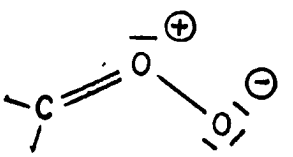


I III II

Whether the ring opening of I involves concerted or stepwise cleavage is unsettled, but that ring opening to the fragments shown does occur is proven by the incorporation of dissolved aldehydes into the secondary ozonide. Structure III, deliberately with only the atomic connectivity shown above, is the crucial intermediate, and is known variously as the Criegee intermediate, carbonyl oxide, or, with hydrogens attached, as peroxymethylene. The Criegee intermediate was proposed to be zwitterionic by Criegee (1957),

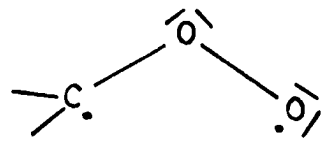


but is sometimes shown as



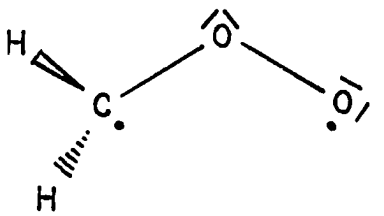
The nature of the Criegee intermediate is of particular interest for ozonolysis in the gas phase, where lower pressure causes the recombined secondary ozonide to be found in only small amounts at long reaction times. The gas phase reaction has been of recent interest due to its important role in photochemical smog generation (Demerjian et al., 1974). The major products of the gas phase reaction of ozone and ethylene are small molecules such as H_2O, CO_2, CO, H_2, H, OH, and traces of larger molecules, e.g., formic acid and methanol (Kuhne et al., 1976).

In their seminal paper, Wadt and Goddard (1975), recognizing that peroxymethylene is isoelectronic to ozone, demonstrated via GVB-CI calculations that peroxymethylene is a singlet biradical,

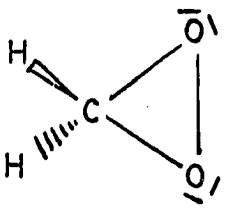


This is exactly analogous to ozone and, in fact, the zwitterions shown above best represent several excited states above 3.5 eV. The biradical nature of the Criegee intermediate had been missed by theorists (Ha et al., 1974) and experimentalists (Criegee, 1957 and Bauld et al., 1968) alike, and has immediate mechanistic consequences. Harding and Goddard (1978) estimated the energy for splitting the primary ozonide to form the Criegee intermediate as about 10 Kcal/mole, leaving the fragments some 40 Kcal/mole below the initial reactants. They calculate the

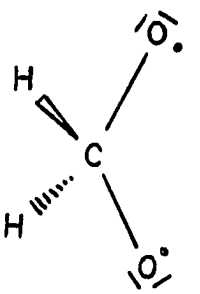
rotational barrier to perpendicular peroxymethylene



as 29 Kcal/mole, so that this rotation is energetically feasible. Similar numerical estimates led Wadt and Goddard to propose that the perpendicular form, now with both unpaired electrons in overlapping orbitals in the COO plane, should undergo ring closure to form dioxirane,



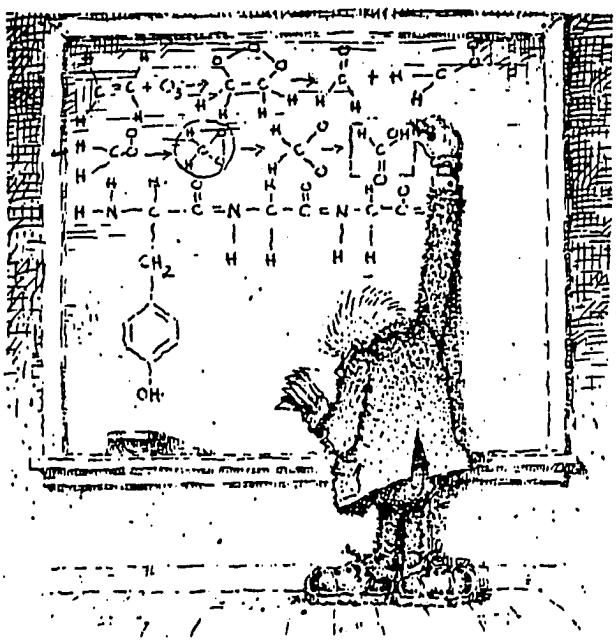
They further propose that this molecule should undergo ring opening to dioxymethane,



which is also a singlet biradical, again by analogy to ozone.

Dioxymethane can undergo a facile 1,2 hydrogen atom shift to formic acid (downhill by 95 Kcal/mole), in a highly vibrationally excited form, which could then decompose into a number of smaller species.

The Wadt-Goddard mechanism for gas phase ozonolysis was confirmed by the discovery of dioxirane in 1977 by two teams performing microwave spectroscopy and mass spectrometry studies of the low temperature gas phase ozonolysis of ethylene. The dramatic discovery of dioxirane was widely reported in such popularizing magazines as Chemistry, Dimensions, and Chemical and Engineering News. The Wadt-Goddard mechanism was even incorporated into a beer ad in the Iowa State Daily:



The actual report of the isolation of dioxirane at temperatures below -100°C came in companion papers. Lovas and Suenram (1977) found a number of microwave transitions attributable to an asymmetric top with a large dipole moment. Martinez, Huie and Herron (1977) confirmed

the H_2CO_2 composition by mass spectrometry. The molecular, as opposed to radical, nature of the species was demonstrated by the ability of these researchers to repeatedly freeze and vaporize the intermediate. Eventually, Suenram and Lovas (1978), by preparing isotopically substituted species, were able to obtain a complete structure for dioxirane.

According to the Wadt-Goddard mechanism, ozonolysis of more complex olefins than ethylene should yield substituted dioxiranes. To date none have been found, which Suenram and Lovas (1978) attribute to unfavorable vapor pressure versus decomposition temperature behavior. Catalan et al. (1980) have presented theoretical geometries for fluorine and methyl substituted dioxirane.

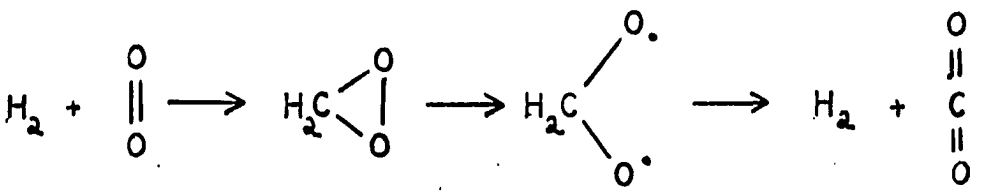
The Wadt-Goddard mechanism leading to final products is the opening of the OO bond of dioxirane to dioxymethane, hydrogen shift to formic acid, which vibrationally decomposes. Martinez, Huie and Herron (1977) found the following final products for the ozonolysis of ethylene, all having the same composition as formic acid: $CO + H_2O$ (67%), $CO_2 + H_2$ (18%), and $CO_2 + 2H$ (9%). Formic acid itself has been detected in the amount of 1% by Kuhne et al. (1976). Qualitative consideration of the electronic structure of dioxymethane (Wadt and Goddard, 1975) shows that the H shift to produce formic acid is highly exoergic. There is likely no barrier to the shift, although there is no concrete evidence to date to support this claim.

The only questionable step in the suggested decomposition of dioxirane at temperatures over -80°C is the ring opening to dioxymethane. Wadt and Goddard (1975) and Harding and Goddard (1978) use assumed

geometries for both dioxirane and dioxymethane, yielding ambiguous energy comparisons. Wadt and Goddard find dioxymethane to be 14 Kcal/mole below the ring state, while Harding and Goddard find dioxymethane to be 11 Kcal/mole above the ring. Karlstrom, Engstrom and Jonsson (1979) investigated the ring opening to dioxymethane, optimizing geometries and locating the transition state. They found dioxymethane to be less than 2 Kcal/mole above the ring, separated by only a 15 Kcal/mole barrier. It is worth noting here that the ring form of ozone lies 28 Kcal/mole above the biradical, open form ground state (Hay and Dunning, 1977). The experimental isolation of dioxirane, the theoretical demonstration of the ring opening, and the final product analysis all firmly establish the Wadt-Goddard mechanism for gas phase ozonolysis.

2. Object of present study

The research described here investigates the entire C_{2v} constrained potential energy surface of the dioxirane/dioxymethane system. Specifically, the reactions considered are



The three reactions shown above all involve extensive configuration mixing, and as a consequence should be treated at the MCSCF level. A

theoretical study of the first reaction, namely the insertion of singlet methylene into singlet oxygen, has never been reported. It should be mentioned that the C_{2v} insertion is not expected to be the most likely reaction route. Methylene usually attacks in a sideways rather than head-on fashion, as for example, in the isoelectronic insertion of singlet methylene into ethylene to produce cyclopropane (Zurawski and Kutzelnigg, 1978). The second reaction, conversion of dioxirane to dioxymethane, is of great significance to the ozonolysis mechanism, as mentioned above. While this conversion has been studied by Karlstrom, Engstrom and Jonsson (1979), their MC-CI results may be biased by obtaining their MCSCF wavefunction at all points with the same ten configurations. These authors extend their probe of the potential surface to the transition state for the final reaction, the C_{2v} dissociation of dioxymethane. The conversion of dioxirane to dioxymethane as well as the C_{2v} dissociation of dioxymethane has been investigated by Cimiraglia, Ha and Gunthard (1982). These authors, too, may have skewed their results by treating the reaction with a multireference CI function using the same six configurations as reference list throughout. In addition, they performed only partial geometry optimization. It should be mentioned that, although $H_2 + CO_2$ is a major product in gas phase ozonolysis, these are formed, according to the Wadt-Goddard (1975) mechanism, on a reactive path leading through formic acid.

All previous investigations on the C_{2v} reactive path connecting methylene and oxygen to carbon dioxide and molecular hydrogen are incomplete in that they suffer either from inadequate configuration

selection procedures or lack of geometry optimization. Therefore, the FORS model has been applied to the reactions, as the predominant configurations may vary greatly from point to point on the potential surface. Full geometry optimization at the FORS level was carried out, and the stable (in C_{2v}) structures and the transition states connecting these relative minima on the potential energy surface were located. The low-lying excited states between dioxirane and dioxymethane are examined.

B. Configurational Description of the Reaction

The subsequent discussion will demonstrate that a Hartree-Fock SCF function is completely incapable of describing the transition states and even some of the stable structures on the C_{2v} reaction coordinate. Since the SCF predictions for the H_2CO_2 system are often wrong, they will not be discussed in any detail. A proper MCSCF methodology is clearly essential for this reactive system.

Ideally, the full valence space type of FORS wavefunction should be used to describe any reactive system as complicated as that of the dioxirane system. The full valence space treats <u>all</u> reactive channels on an equal footing, so that the transformation of carbon from singly bonding two hydrogens to doubly bonding two oxygens is handled in a single consistent and unbiased fashion. A quick check shows that, even with the simplification due to C_{2v} symmetry, the full valence space for a $^1A_1$ state contains 42,816 SAAPs. As this is too large a number, the reaction sequence is divided into two parts: first, the insertion of methylene into oxygen to form dioxirane and subsequently dioxymethane

and, second, the dissociation of dioxymethane. Each part of the reaction is treated with a reduced FRS appropriate to that particular reaction, with the two treatments overlapping around the dioxymethane portion of the potential energy surface.

1. Formation of dioxirane and dioxymethane

During the insertion of methylene into oxygen to form first dioxirane and then dioxymethane, the oxygens vary from doubly to singly to not bound. At the same time, new CO bonds are formed. However, the CH bonds are not broken, and one lone pair possessing largely 2s character persists on each oxygen. These four orbitals together with the oxygen and carbon 1s inner shells are taken as the core orbital space. In C_{2v} symmetry notation, it consists of three a_1, one b_1, and two b_2 orbitals as shown in Table 5.1. Here the x axis is perpendicular to the COO plane, the y axis is parallel to the OO bond, and the z axis lies on the twofold rotation axis. This core space is occupied by 6 inner shell and 8 valence electrons.

The remaining 8 orbitals form the reactive orbital space which can be characterized from the atomic and from the molecular points of view. The filled σ and π, half-filled π^*, and empty σ^* shells of molecular oxygen which arise predominantly from the 2p O atom orbitals allow for correct dissociation of O_2 to 2O, and permit the oxygen atoms to combine with the methylene. On the carbon atom, only the sp hybrid lying in the CH_2 plane, and the p orbital perpendicular to the HCH plane (and lying in the COO plane) are involved in the reduced FRS of a methylene whose CH bonds are kept doubly occupied. The combined set of these

Table 5.1. Symmetry labels for the orbitals during insertion

Symmetry Label	Orbital Type	O_2 or CH_2	Dioxirane/ Dioxymethane
$1a_1$	core	O 1s + O 1s	O 1s + O 1s
$2a_1$	core	C 1s	C 1s
$3a_1$	core	O ℓp + O ℓp	O ℓp + O ℓp
$4a_1$	core	CH + CH	CH + CH
$5a_1$	reactive	OO σ	OO σ
$6a_1$	reactive	C ℓp	CO + CO
$7a_1$	reactive	OO π	CO^* + CO^*
$1a_2$	reactive	OO π^*	OO π^*
$1b_1$	core	CH - CH	CH - CH
$2b_1$	reactive	OO π	OO π
$1b_2$	core	O 1s - O 1s	O 1s - O 1s
$2b_2$	core	O ℓp - O ℓp	O ℓp - O ℓp
$3b_2$	reactive	OO π^*	CO - CO
$4b_2$	reactive	OO σ^*	OO σ^*
$5b_2$	reactive	C p	CO^* - CO^*

orbitals for methylene and oxygen consists of three a_1, one a_2, one b_1, and three b_2 MOs, as shown in Table 5.1.

For the molecule dioxirane, the two CO sigma bonds must be left-right correlated to account for their formation as the methylene approaches the oxygen. The OO sigma bond must also be left-right correlated for that bond to be broken in forming dioxymethane. Dioxirane also possesses filled OO π and π^* oribtals (leading to no net OO π bonding) which must be retained in the reactive orbital set for two reasons. First, these orbitals correlate with the out of plane π and π^* orbitals of O_2, whose in plane counterparts participate in bonding and antibonding to the approaching CH_2 group. A consistent treatment of the isolated oxygen requires both π and π^* orbitals be in the reactive set. A second reason is that these π and π^* orbitals are single occupied for dioxymethane. The very long OO distance in dioxymethane removes the energy penalty for occupying the OO sigma antibond, so the repulsion of the out of plane CH bond pairs causes only two electrons to populate the π and π^* shells, with four electrons in the σ and σ^*. As shown in Table 5.1, the reactive orbitals for dioxirane and dioxymethane have the same symmetry labels as those of the separated methylene and oxygen. Therefore, the reactive orbital space for the insertion of methylene into oxygen to form dioxirane and the subsequent conversion to dioxymethane is chosen to consist of three a_1, one a_2, one b_1, and three b_2 MOs, correlating ten valence electrons.

Clearly, the electronic state of dioxirane is $^1A_1$ (one can draw a Lewis structure for it), with the SCF configuration

$$\text{core } (5a_1)^2(6a_1)^2(1a_2)^2(2b_1)^2(3b_2)^2 \quad .$$

The $^1A_1$ FRS for the orbital space just described contains 320 configurations, a number which is readily handled by the macro-iterative optimization procedure of Dombek and Ruedenberg described in Chapter II. The lowest potential surface for dioxirane must connect with the lowest $^1A_1$ states of dioxymethane and the separated methylene and oxygen. For dioxymethane, this state is a two configuration wavefunction:

$$\text{core } (5a_1)^2(6a_1)^2(3b_2)^2(4b_2)^2[(1a_2)^2 - (2b_1)^2] \quad ,$$

so that on the average the oxygen π and π^* orbitals are singly occupied. The biradical dioxymethane must also possess a low-lying triplet state, namely

$$\text{core } (5a_1)^2(6a_1)^2(3b_2)^2(4b_2)^2(1a_2)^1(1b_1)^1 \quad .$$

The separated methylene and oxygen must be in a combined $^1A_1$ state to form the ground state of dioxirane. The ground states of both molecules are triplets, which can, of course, couple to a combined singlet. However, labeled in the combined C_{2v} coordinate system, triplet methylene is in a $^3B_2$ state and triplet oxygen is $^3B_1$, so that the combined spatial symmetry is A_2, not A_1. Therefore, the triplets in a C_{2v} path correlate with an excited state of dioxirane. Of course, singlet oxygen and triplet methane, or triplet oxygen and singlet methylene cannot react adiabatically to give the singlet dioxirane. Singlet methylene and singlet oxygen can react to form the ground state of

dioxirane, as singlet methylene is in a $^1A_1$ state, and in C_{2v} symmetry the degenerate singlet Δ state of O_2 resolves into a $^1A_1$ and a $^1B_1$ component. Consequently, at long separations, one expects the two configurations

$$\text{core } (5a_1)^2(6a_1)^2(7a_1)^2(2b_1)^2[(3b_2)^2 - (1a_2)^2]$$

to be predominant. This function is the product of the SCF function for methylene and the open shell SCF function for the $^1A_1$ component of singlet delta oxygen. Note that in lower symmetry than C_{2v}, i.e., sideways attack of methylene, the reaction of triplet methylene and triplet oxygen to form singlet dioxirane is not forbidden on orbital symmetry grounds.

The two methylene reaction orbitals generate a two configuration FRS for $^1A_1$ methylene. The six oxygen reaction orbitals lead to a FRS containing 18 SAAPs for the $^1\Delta_g$ state of oxygen. The product wavefunction for the two isolated species thus contains 36 SAAPs (out of the full 320) with nonvanishing coefficients.

The dominant configurations for the description of separated singlet methylene and singlet oxygen, dioxirane, and dioxymethane given in the preceding paragraph are different. These changes in the character of the lowest $^1A_1$ state must be due to avoided crossings with the second $^1A_1$ surface. Because of these avoided crossings, methylene plus oxygen, dioxirane, and dioxymethane are each separated by a barrier.

2. Dissociation of dioxymethane

During the course of the C_{2v} dissociation of dioxymethane to molecular hydrogen and carbon dioxide, the CH bonds are broken, an HH bond forms, and CO pi bonds are formed. However, the CO sigma bonds formed during the methylene insertion remain as do the O lone pairs, one on each oxygen. These four persistent orbitals are now kept in core together with 1s inner shells.

To correctly treat the CH bond breakage, these bonds must be left-right correlated by including orbitals arising from the hydrogen 1s and two carbon sp hybrids. Concomitant with the hydrogen separation, the two hydrogen 1s orbitals form the bond and correlating antibond of molecular hydrogen, giving a two configuration wavefunction for an isolated hydrogen molecule.

The carbon dioxide fragment must form new pi bonds. For linear carbon dioxide, the pi bonding, nonbonding, and antibonding orbitals arise from the two p_π orbitals on each oxygen and the carbon. The two sp hybrids on carbon, following the splitting off of the hydrogens will become p_π orbitals as the CO_2 linearizes. The other two sp type sigma hybrids on carbon are involved in bonding to a hybrid orbital on oxygen forming the persistent CO sigma bond orbitals which are kept in the core as is the one lone pair orbital on each oxygen. The remaining two orbitals on each oxygen form, in dioxymethane, the nearly doubly occupied OO sigma bond and antibond, and the almost singly occupied OO pi bond and antibond. As the dissociation proceeds and the CO_2 becomes linear, these pi orbitals become oxygen p_π orbitals out of the original

COO plane, while the sigma orbitals lose their s character, becoming p_π orbitals in the original COO plane. Thus, the carbon and oxygen orbitals needed to describe dioxymethane as the hydrogens are splitting off become the full set of pi orbitals for linear CO_2. For the $^1\Sigma_g^+$ ground state of CO_2, the FRS built from the full π orbital space contains 33 configurations. Note that this orbital space cannot properly dissociate CO_2, but does correctly describe the <u>bending</u> of CO_2.

The symmetry labels of these orbitals are shown in Table 5.2. It is apparent that the orbitals used to describe dioxymethane, and the orbitals needed to describe separated hydrogen and carbon dioxide belong to the same symmetries: there are three a_1, one a_2, three b_1, and one b_2 MOs correlating 10 valence electrons. The core orbital space consists of four a_1 and three b_2 orbitals, also shown in Table 5.2, containing six inner shell and eight valence electrons. The $^1A_1$ FRS again consists of 320 configurations. It contains all 2 x 33 SAAPs of the dissociated states of H_2 and CO_2, so that after dissociation 66 of the 320 SAAPs have nonvanishing coefficients.

To clarify the difference in the choice of core, reactive, and virtual MOs for the insertion and ring opening reactions, and that for the dissociation reaction, Table 5.3 compares all these MOs for dioxymethane, where the two FRS treatments overlap. It should be noted that those orbitals (CO and CH bonds and antibonds) which are switched between core, reactive, and virtual spaces carry a different ordering number in their symmetry labels in the two cases. It is obvious that part of the 320 SAAPs are the same and part different in the two cases.

Table 5.2. Symmetry labels for dioxymethane dissociation

Symmetry Orbital	Orbital Type	Dioxymethane	CO_2 or H_2
$1a_1$	core	O 1s + O 1s	O 1s + O 1s
$2a_1$	core	C 1s	C 1s
$3a_1$	core	O ℓp + O ℓp	O ℓp + O ℓp
$4a_1$	core	CO + CO	CO + CO
$5a_1$	reactive	OO σ	CO_2 π_u
$6a_1$	reactive	CH + CH	H_2 σ
$7a_1$	reactive	CH^* + CH^*	CO_2 π_u^*
$1a_2$	reactive	OO π^*	CO_2 π_g
$1b_1$	reactive	CH - CH	CO_2 π_u
$2b_1$	reactive	OO π	CO_2 π_u^*
$3b_1$	reactive	CH^* - CH^*	H_2 σ^*
$1b_2$	core	O 1s - O 1s	O 1s - O 1s
$2b_2$	core	O ℓp - O ℓp	O ℓp - O ℓp
$3b_2$	core	CO - CO	CO - CO
$4b_2$	reactive	OO σ	CO_2 π_g

The wavefunction for dioxymethane consists predominantly of two configurations

$$\text{core } (5a_1)^2(6a_1)^2(1b_1)^2(4b_2)^2[(1a_2)^2 - (2b_1)^2] \quad .$$

This is exactly the same as the two predominant configurations given above in the discussion of the insertion FRS, as may be seen by writing out explicitly the core portion in each case and interpreting the symmetry labels according to Table 5.3. The predominant term in the wavefunction for separated hydrogen and carbon dioxide is the product of their SCF functions,

$$\text{core } (5a_1)^2(6a_1)^2(1a_2)^2(1b_1)^2(4b_2)^2$$

in C_{2v} notation. As this is formally the same as the first term of the basic dioxymethane function, the changes in electronic structure during the dissociation of dioxymethane are simpler than during the methylene insertion, or conversion of dioxirane to dioxymethane. An MCSCF treatment of the predominantly two configuration dioxymethane is a necessity, but the wavefunction actually simplifies during the course of the dissociation to a single predominant term. The four electrons in the CH bonds form the H_2 bond and the out of plane carbon dioxide bond. Simultaneously an in plane pi bond forms.

C. Calculational Details

1. Selection of reaction coordinate

Much chemical significance is attributed to stationary points on molecular potential energy surfaces, where the first derivatives of the

Table 5.3. Comparison of the two FRS orbital sets

Insertion and Ring Opening		Common		Dissociation
		Core Orbital Space		
	$1a_1$	O 1s + O 1s	$1a_1$	
	$1b_2$	O 1s - O 1s	$1b_2$	
	$2a_1$	C 1s	$2a_1$	
	$3a_1$	O ℓp + O ℓp	$3a_1$	
	$2b_2$	O ℓp - O ℓp	$2b_2$	
CH + CH	$4a_1$		$4a_1$	CO + CO
CH - CH	$1b_1$		$3b_2$	CO - CO
		Reactive Orbital Space		
	$5a_1$	OO σ	$5a_1$	
	$4b_2$	OO σ^*	$4b_2$	
	$2b_1$	OO π	$1b_1$	
	$1a_2$	OO π^*	$1a_2$	
CO + CO	$6a_1$		$6a_1$	CH + CH
$CO^* + CO^*$	$7a_1$		$7a_1$	$CH^* + CH^*$
CO - CO	$3b_2$		$2b_1$	CH - CH
$CO^* - CO^*$	$5b_2$		$3b_1$	$CH^* - CH^*$
		Virtual Orbital Space		
$CH^* + CH^*$	$8a_1$		$8a_1$	$CO^* + CO^*$
$CH^* - CH^*$	$3b_1$		$5b_2$	$CO^* - CO^*$

total energy with respect to all nuclear coordinates vanish. Stable molecules are at an energy minimum with respect to all nuclear coordinates, so that all second derivatives (force constants) are positive and, hence, all molecular frequencies are real. The location of such stable structures is straightforward; the energy must be minimized with respect to all nuclear coordinates.

A transition state is a saddle point on the potential surface with one and only one negative second derivative and, of course, vanishing first derivatives. Starting out in the direction of negative curvature and proceeding downhill leads to products or back to reactants. For this reason, this direction is termed the reaction coordinate, and for motion along the reaction coordinate, the energy is maximal at the transition state. At the transition state, the reaction coordinate corresponds to a normal mode, whose force constant (proportional to the second derivative of the energy) is negative. The formal frequency (proportional to the square root of the force constant) is thus imaginary, implying that motion along the reaction coordinate falls off to reactants or products without return ("vibration") to the transition state.

The location of a transition state is a more difficult problem than that of finding stable structures since unconstrained minimization of the the energy causes collapse into either the reactants or products. One possibility is to minimize the sum of the squares of all the first derivatives of the energy with respect to nuclear displacement. This length of the gradient vector goes to zero at transition states as well

as at stable structures. Recently the analytic calculation of energy gradients has become a very powerful tool for the investigation of potential surfaces (Flanigan et al., 1977). Unfortunately, most current programs using gradients are limited to SCF type wavefunctions. Dupuis (1981) has recently extended the gradient formalism to MCSCF wavefunctions, but programs to calculate MCSCF gradients are not yet common.

Since the ALIS program system does not currently permit the analytic calculation of gradients, a more cumbersome "brute force" method is used in the present study. Were it possible to know the reaction coordinate, the minimization of the total energy in all directions save the reaction coordinate and maximization of the energy along the reaction coordinate would yield the transition state. Therefore, if it is possible to select a "good" reaction coordinate, then the following procedure may be applied. At each value of this reaction coordinate, the total energy is minimized with respect to all other coordinates. A plot of the energy versus the chosen reaction coordinate after optimization of the other coordinates possesses minima, which are the stable structures, and maxima which are the transition states. The reliability of this approach compared to the accurate gradient method is related to the ability to preselect a reaction coordinate that enters the transition state in a direction close to the eigenvector corresponding to the imaginary frequency. This approach is also inefficient in computer time, as energy evaluations are required at more geometries than with an analytic gradient approach.

In applying this method to the H_2CO_2 surface, a reaction coordinate had to be chosen. Clearly the distance from the methylene to the oxygen molecule governs the progress of the insertion reaction. Therefore, the distance from the carbon atom to the midpoint of the OO line is chosen as the reaction coordinate for the methylene insertion to form dioxirane. This distance is also taken as the reaction coordinate for the ring opening to dioxymethane, as the ring strain engendered by pushing the methylene group even closer to the OO bond than in the near equilateral triangular dioxirane will cause the OO bond to be broken. Similarly, for the dissociation of dioxymethane to carbon dioxide and hydrogen, the reaction coordinate is defined as the distance from the carbon atom to the center of the HH line.

Assuming a C_{2v} symmetry with the HCH plane orthogonal to the OCO plane, the geometry of the H_2CO_2 system is described by four coordinates. Thus, three coordinates must be optimized for each value of the above defined reaction coordinate. During the insertion and ring opening reactions, these coordinates can be taken as the CH distance, the HCH angle, and the CO distance. As CH bond distances vary very little, the optimizations can be reduced to two coordinates by assuming the CH bond length is 1.09 Å. The CO bond distance is a natural coordinate to optimize around the dioxirane/dioxymethane portion of the surface. At long distances of the methylene to oxygen, the optimization of the OO bond length in place of the CO distance is more natural. Optimal values for both the OO and CO distances are given below as a function of reaction coordinate but, of course, only one of these variables is independent.

For the dioxymethane dissociation, no similar reduction of the optimization problem can be made. The coordinates which are optimized are the HH distance, the CO distance, and the OCO angle, at each value of the carbon to H_2 reaction coordinate. The CO bonds shorten considerably as they go from single to double bonds, so they cannot be fixed as the CH bonds were for the insertion reaction.

The geometry optimization for a particular value for the reaction coordinate is done by quadratic fit to energies calculated at several trial values of the coordinates being optimized. For optimization of two or three coordinates, these fits require energy calculations at six or ten points, respectively. Least mean square fits are made if more than the minimum number of points are available. If the trial points are well-selected, the optimal geometry is given by the minimum of the fitted quadratic. If the initial guess of the molecular structure is less good or the surface is not very quadratic, more energy evaluations at better chosen trial geometries are required. A typical optimization requires ten or fifteen energy evaluations to optimize two or three independent coordinates.

2. Basis sets

In view of the large number of different geometries for which the wavefunction and energy must be obtained, the size of the basis set must be kept relatively small. The basis set used for the geometry optimization is (8s,4p/3s,2p) on the heavy atoms, and (3s/2s) (unscaled) on hydrogen. This totals 33 QBOs, which is manageable for the large number of calculations needed. It is well-known, however, that bond

lengths in the isoelectronic molecule ozone are not predicted well unless polarization functions are included. Hay and Dunning (1977) found OO bonds to be 0.1 Å too long when polarization functions are not used in ozone, but that this overestimate is easily corrected by including a single set of d functions. Unfortunately, adding d orbitals on the heavy atoms to the basis used here, even with the 3s component transformed away, would increase its size to 48 QBOs, making the geometry optimizations too costly. Therefore, the smaller unpolarized basis is used for geometry optimization, even though this probably systematically overestimates CO and OO bond lengths by around 0.1 Å. A larger basis, including also polarization functions on the hydrogen, is used to obtain the energy at the various stable geometries and transition states, once these are located with the smaller basis. This larger basis, which is (12s,6p,1d/3s,2p,1d) on the heavy atoms and (4s,1p/2s,1p) (unscaled) on hydrogen contains 55 QBOs and should give more accurate estimates of the reaction energetics.

The primitive orbital sets are the even-tempered Gaussian bases given by Schmidt and Ruedenberg (1979) contracted in Raffenetti style as described in Chapter II. Polarization exponents are taken from Dunning and Hay (1977).

D. Results and Discussion

1. Formation of dioxirane and dioxymethane

For each value of the methylene to oxygen reaction coordinate, both remaining geometric parameters have been optimized. The geometry

optimizations were conducted at the FORS rather than SCF level. Optimization of the orbitals in the 320 SAAP reaction space used to describe the insertion reaction typically required eight to fifteen configurations in the MCSCF calculations. The largest number of SAAPs in the MCSCF orbital optimizations is required around the two transition states, and for dioxymethane, where two dominant configurations must be correlated. The configurations which dominate the expansion of the FORS wavefunction vary greatly as the reaction coordinate is changed, and even vary somewhat during the optimization of the other geometric coordinates for a fixed reaction coordinate. The MCSCF configurations used for the orbital optimizations were, for each geometry, the predominant configurations based on natural reaction orbitals in the final FORS CI wavefunction at that particular geometry. For the methylene insertion and the subsequent conversion of dioxirane to dioxymethane, about 30 different SAAPs in total were used for the MCSCF optimizations at one or another of the geometries considered. The final wavefunctions were in all cases 320 SAAP CI functions based on the optimal orbitals.

The energies of the C_{2v} H_2CO_2 system during the insertion of methylene to form dioxirane, and the ring opening to form dioxymethane, are shown in Fig. 5.1. The optimal geometric parameters for each value of the reaction coordinate are given in Table 5.4 and are displayed in Figs. 5.2, 5.3 and 5.4. On the basis of these graphs, a number of geometries were interpolated between the optimized ones and the energies were rigorously optimized at these interpolated geometries to flesh out the energy curve in Fig. 5.1. These interpolated geometries are also

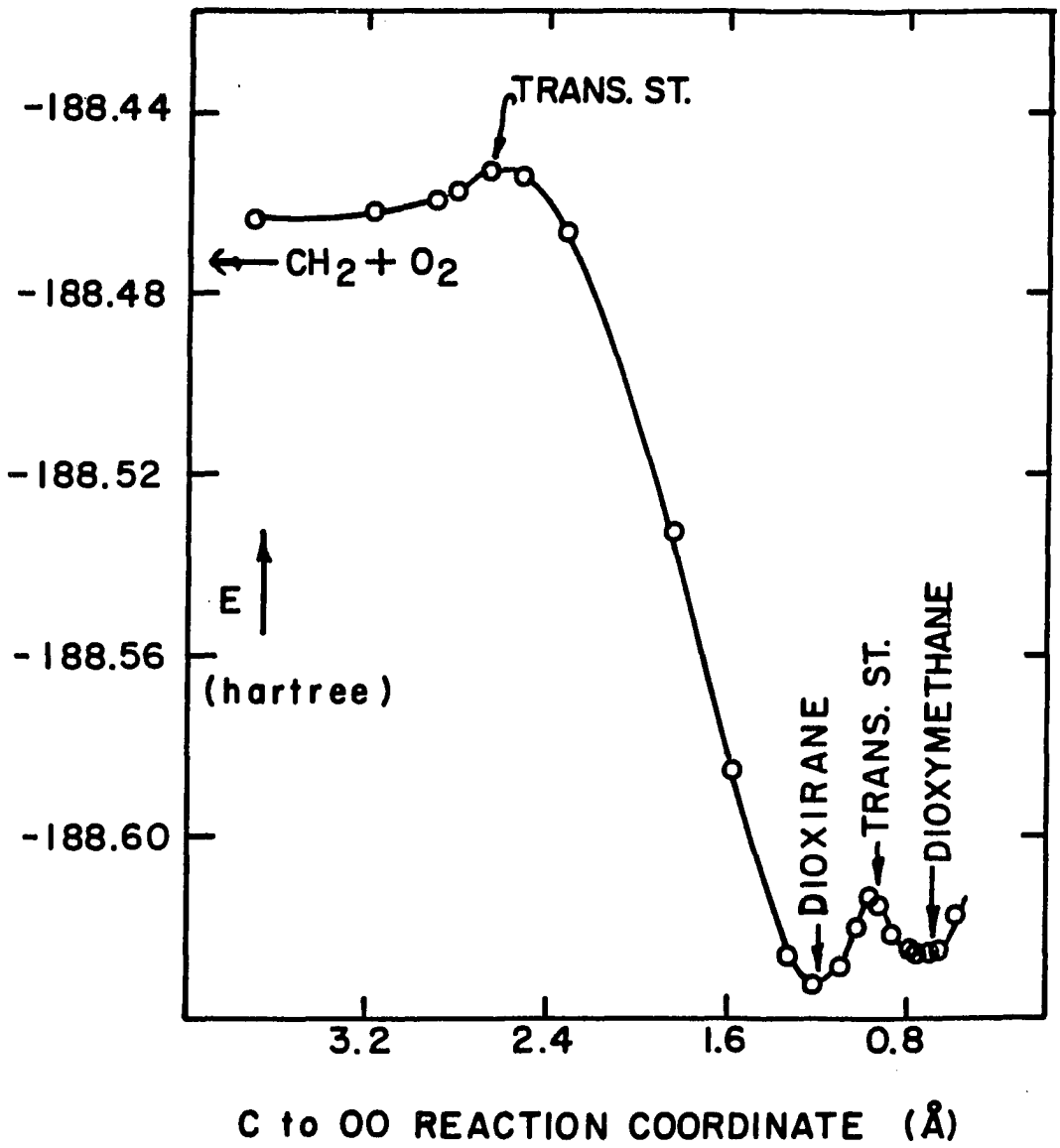


Figure 5.1. Total energy during methylene insertion and ring opening of dioxirane

Table 5.4. Geometries and energies for formation of dioxirane and dioxymethane[a]

Reaction Coordinate (Å)	OO Bond (Å)	CO Bond (Å)	HCH Angle (degrees)	FORS Energy (hartrees)	
infinite	1.334	4.547	105.7	-188.424014	$CH_2 + O_2$
3.704[b]	1.334	3.764	105.7	-188.424292	
3.175[b]	1.334	3.244	105.7	-188.423173	
2.910[b]	1.334	1.986	105.7	-188.420555	
2.805[b]	1.334	2.883	105.7	-188.418625	
2.646	1.336	2.729	104.9	-188.413704	trans. st.
2.514	1.728	2.658	132.9	-188.414786	
2.328	1.650	2.470	136.4	-188.427027	
1.852	1.580	2.014	137.7	-188.493404	
1.588	1.578	1.773	132.4	-188.545905	
1.323[b]	1.640	1.557	122.0	-188.587008	
1.217	1.679	1.479	118.8	-188.594272	dioxirane
1.111[b]	1.799	1.430	114.0	-188.590101	
1.032[b]	1.958	1.422	112.0	-188.581172	
0.979	2.041	1.414	112.2	-188.575312	trans. st.
0.942	2.348	1.505	115.1	-188.576436	
0.873	2.384	1.478	113.6	-188.582865	
0.794	2.405	1.441	112.5	-188.586668	
0.751[b]	2.418	1.424	111.2	-188.587758	dioxymethane
0.714	2.456	1.421	110.0	-188.587318	
0.688[b]	2.466	1.412	108.5	-188.586668	
0.582	2.540	1.397	106.7	-188.579338	

[a]Stationary points indicated in final column. In all cases, CH distance is 1.09 Å.

[b]These geometries are interpolated between the optimal ones.

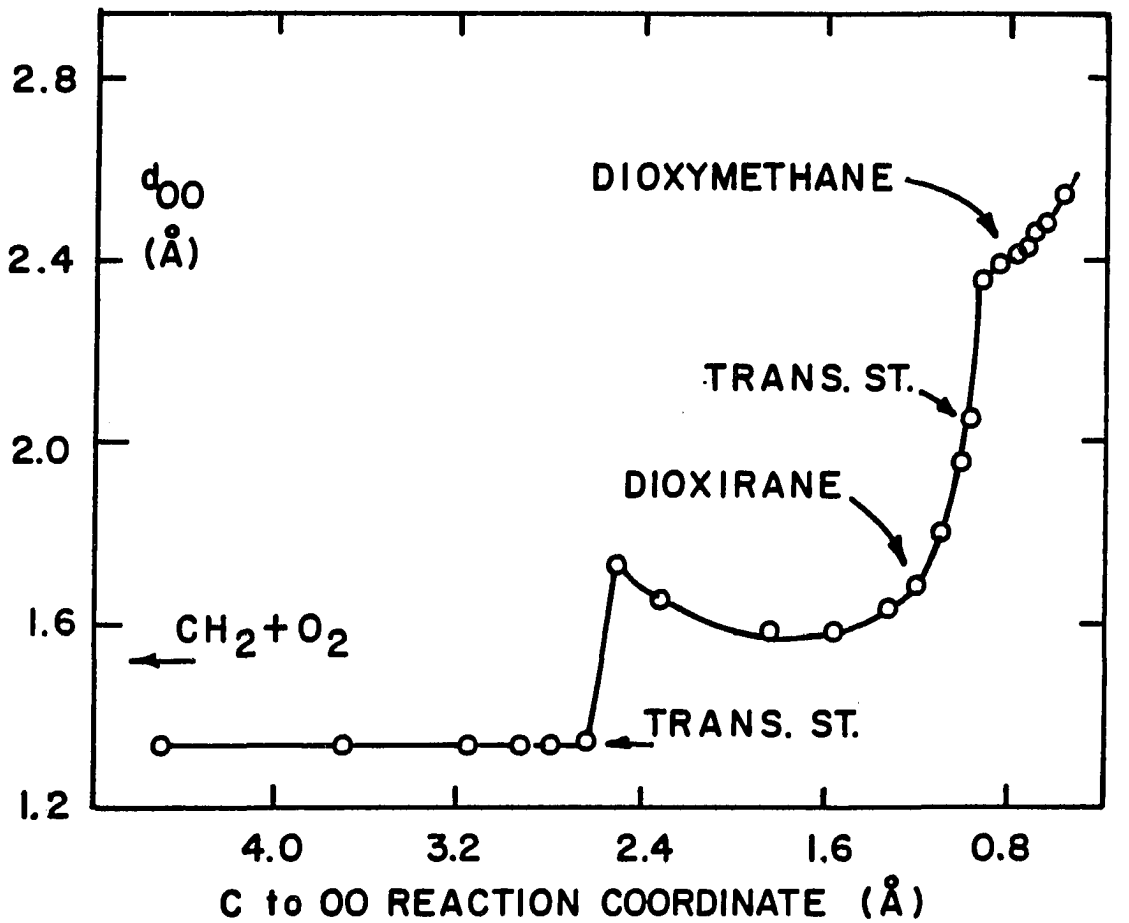


Figure 5.2. Variation of oxygen-oxygen distance during methylene insertion and dioxirane ring opening

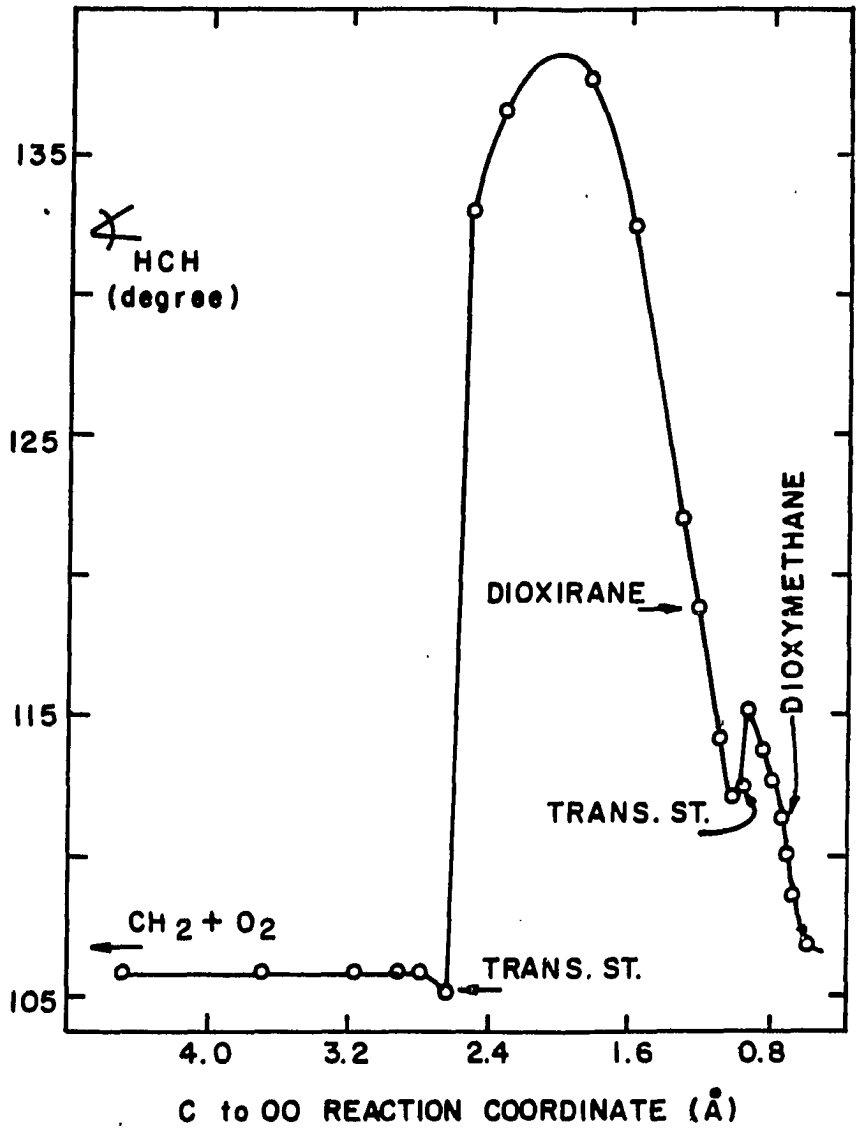


Figure 5.3. Variation of HCH angle during methylene insertion and dioxirane ring opening

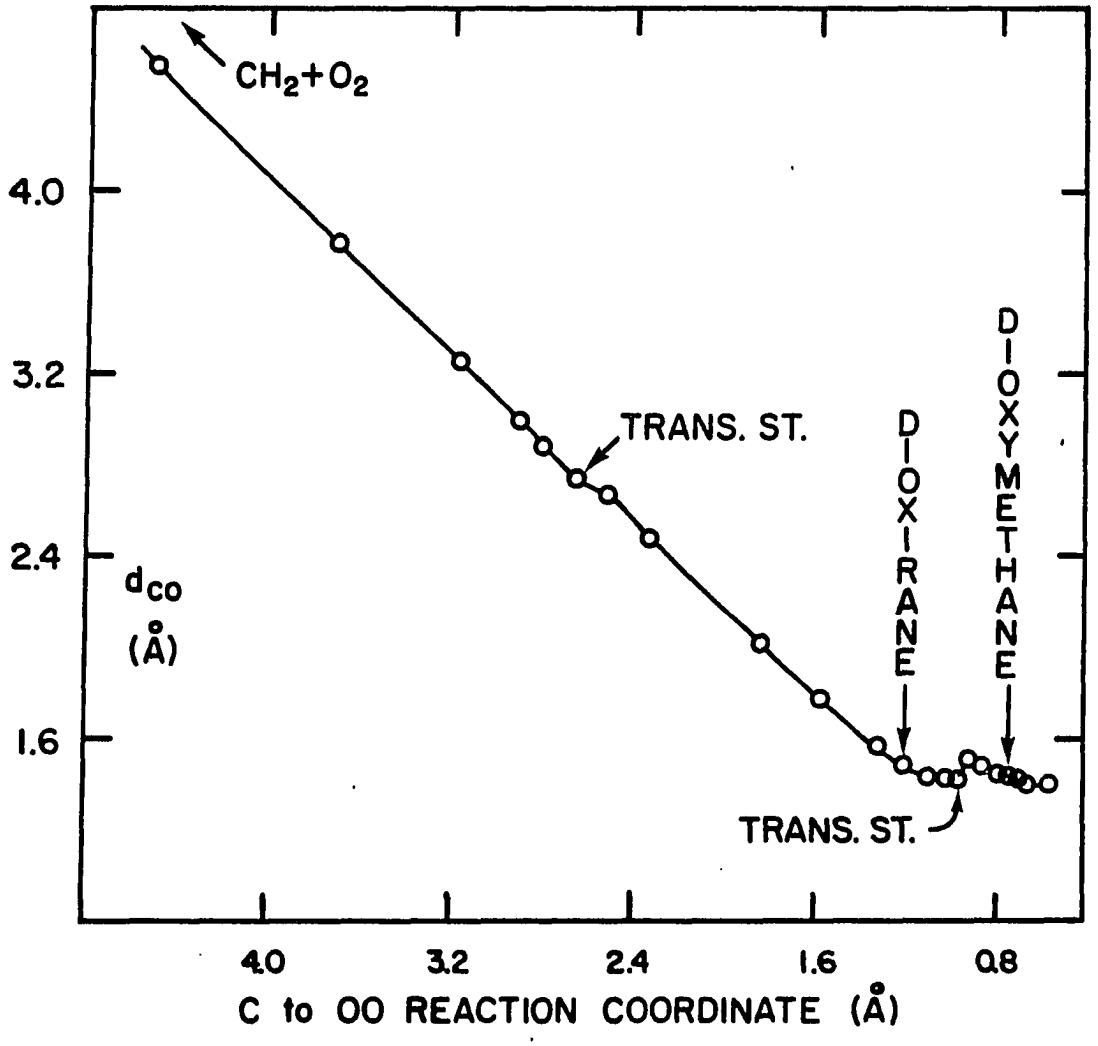


Figure 5.4. Variation of carbon-oxygen distance during methylene insertion and dioxirane ring opening

given in the various figures and table. The transition states and stable structures are the maxima and minima on the energy curve shown in Fig. 5.1.

Inspection of Figs. 5.2, 5.3 and 5.4 shows that there are sudden changes in the geometric parameters as the system crosses the two barriers. This is because the barriers originate in avoided crossings of the two lowest $^1A_1$ surfaces, so that on either side of the barrier on the ground state surface, the electronic structure which determines the geometry is quite different.

The OO bond distances in Fig. 5.2 are quite revealing. The double bond of molecular oxygen is lost just as the barrier is crossed, which is reflected in the sudden jump in the OO distance to a value characteristic of single bonding. This bond is subjected to increasing strain as the system passes through dioxirane, lengthening and finally rupturing, as the OO distance increases to 2.4 Å as the barrier to ring opening is surmounted.

There are similar sudden changes in the HCH angle, shown in Fig. 5.3. Just after crossing the insertion barrier, the HCH angle increases by thirty-five degrees. This is presumably due to the lone pair of singlet methylene being drawn to the incipient CO bonds. The reduced electron repulsion due to lower electron density near the carbon would cause the HCH angle to open. This explanation is consistent with the sharp change to OO single bonding noted above, indicating the oxygen atoms begin bonding to the approaching carbon right as the barrier is crossed. The HCH angle then generally decreases as the OCO angle

increases. This decrease is a "conservation of angles" around the tetracoordinate carbon as the OCO angle opens. The decreasing trend in the HCH angle is interrupted at the conversion barrier. This three degree increase is rather insignificant compared to the loss of the OO bond, and the downward trend in the HCH angle continues as the OCO angle widens still further.

Figure 5.4 shows the CO distance. This coordinate is not independent from the OO distance, and the slight jumps in the CO distance are merely a reflection of the sudden changes from doubly to singly to not bound oxygens, seen more clearly in Fig. 5.2.

2. Dissociation of dioxymethane

The three independent coordinates of the H_2CO_2 system during the dissociation were optimized for each value of the carbon to hydrogen midpoint reaction coordinate. These optimizations were carried out in the 320 SAAP full reaction space used to describe the dissociation. Typically, seven to ten SAAPs were used in the final MCSCF orbital optimization of the full reaction space. The leading SAAPs are little changed in proceeding from dioxymethane to hydrogen and carbon dioxide. The major shift is the increase in importance of one of the two predominant dioxymethane SAAPs at the expense of the other to become the single predominant configuration for H_2 and CO_2.

The energies during the dissociation reaction are plotted versus reaction coordinate in Fig. 5.5. Dioxymethane is separated in C_{2v} symmetry from the dissociation products by a significant barrier. The CH, HH and CO distances and the OCO bond angle during the course of the

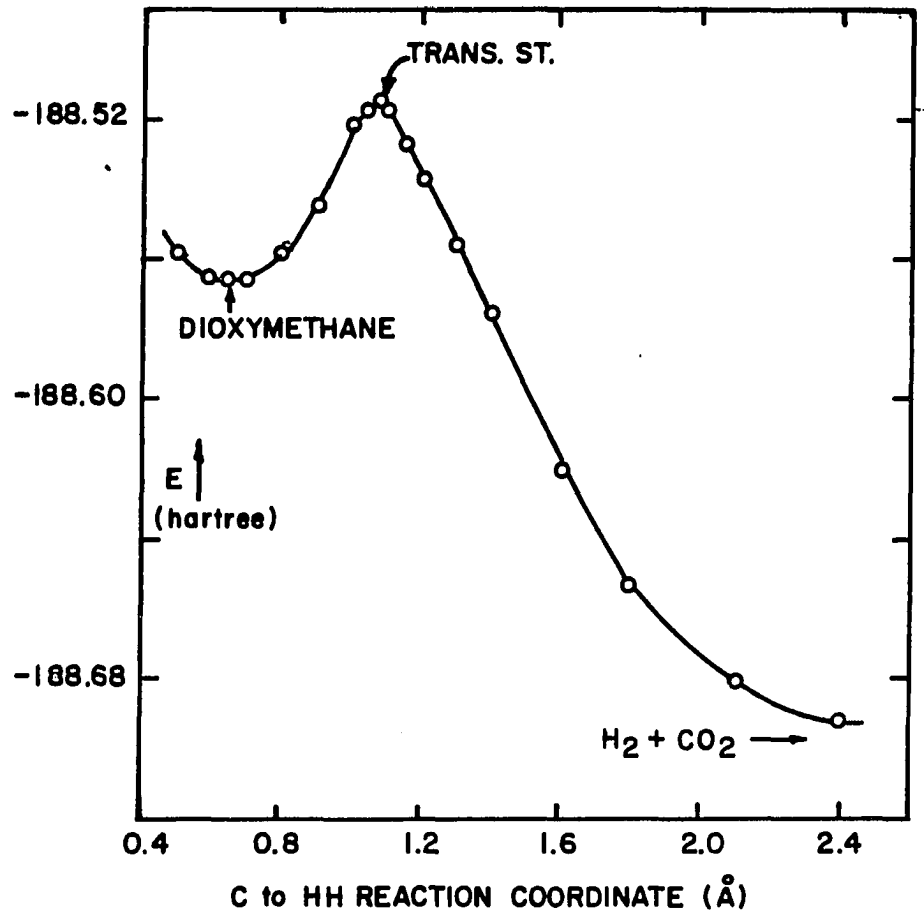


Figure 5.5. Total energy during dissociation of dioxymethane

dissociation are given in Table 5.5. These values are displayed in Figs. 5.6, 5.7, 5.8 and 5.9.

The various geometric parameters shown in Figs. 5.6 through 5.9 are much simpler than those found for the insertion reaction. In keeping with the gradual change in the FORS wavefunction during the dissociation, the various bond lengths and angles vary smoothly from their value in dioxymethane to those in the dissociation products. It is apparent from the figures that the dissociation reaction is essentially complete by the time the hydrogens are 2.4 Å away from the carbon dioxide.

Comparison of Table 5.5 with Table 5.4 shows that the calculated geometries for dioxymethane are slightly different because of the change in the configurations from those needed to describe the formation of dioxymethane to those needed for its dissociation. These geometry differences are, however, small. This dichotomy would not result, of course, if the full valence space were used to treat the entire reaction path. The optimal CH bond lengths found for the early phases of the dioxymethane dissociation are in the range 1.07 Å to 1.10 Å. Thus, the assumption of 1.09 Å for the CH bond length throughout the insertion reaction is justified.

3. Accuracy of calculated geometries

The calculated geometries for the various stationary points on the H_2CO_2 surface were given in Tables 5.3 and 5.4. In the case of the stable molecules, comparison with the experimentally known geometries is possible. For singlet oxygen, the FORS bond length is calculated to be 1.334 Å, while the experimental length is 1.216 Å (Huber and

Table 5.5. Geometries and energies for the dissociation of dioxymethane[a]

Reaction Coordinate (Å)	CH Bond (Å)	HH Bond (Å)	CO Bond (Å)	OCO Angle (degrees)	FORS Energy (hartrees)	
0.500	1.070	0.892	1.417	116.6	-188.558936	
0.600[b]	1.082	1.800	1.405	118.0	-188.565742	
0.650[b]	1.088	1.745	1.402	119.0	-188.566639	dioxymethane
0.700	1.105	1.710	1.394	120.0	-188.565906	
0.800[b]	1.138	1.620	1.380	122.0	-188.558842	
0.900	1.179	1.523	1.394	125.4	-188.544962	
1.000	1.214	1.377	1.330	131.3	-188.522632	
1.050[b]	1.222	1.250	1.310	135.0	-188.518517	
1.081[b]	1.226	1.155	1.296	138.0	-188.516626	trans. st.
1.100[b]	1.225	1.080	1.285	139.0	-188.517908	
1.150	1.238	0.916	1.258	142.9	-188.527790	
1.200[b]	1.266	0.808	1.226	147.5	-188.537083	
1.300[b]	1.357	0.780	1.200	152.0	-188.556565	
1.400	1.452	0.770	1.190	155.0	-188.576254	
1.600[b]	1.645	0.765	1.183	162.0	-188.621415	
1.800	1.840	0.763	1.175	168.0	-188.653772	
2.100[b]	2.134	0.761	1.174	175.0	-188.681189	
2.400[b]	2.430	0.761	1.174	178.0	-188.693096	
infinite	infinite	0.761	1.174	180.0	-188.700398	$H_2 + CO_2$

[a]Stationary points indicated in final column.

[b]Geometry interpolated between optimized values.

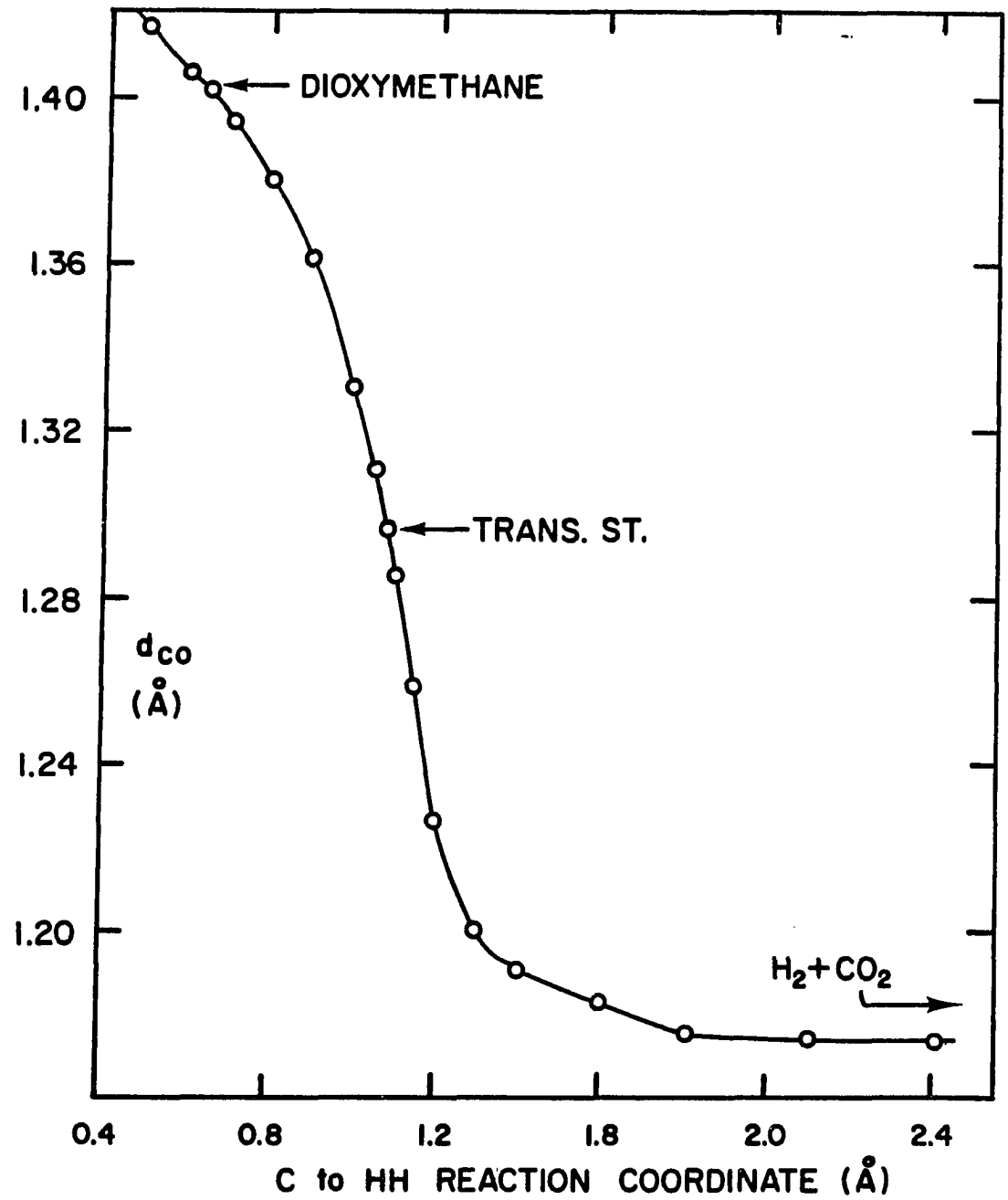


Figure 5.6. Variation of carbon-oxygen distance during dioxymethane dissociation

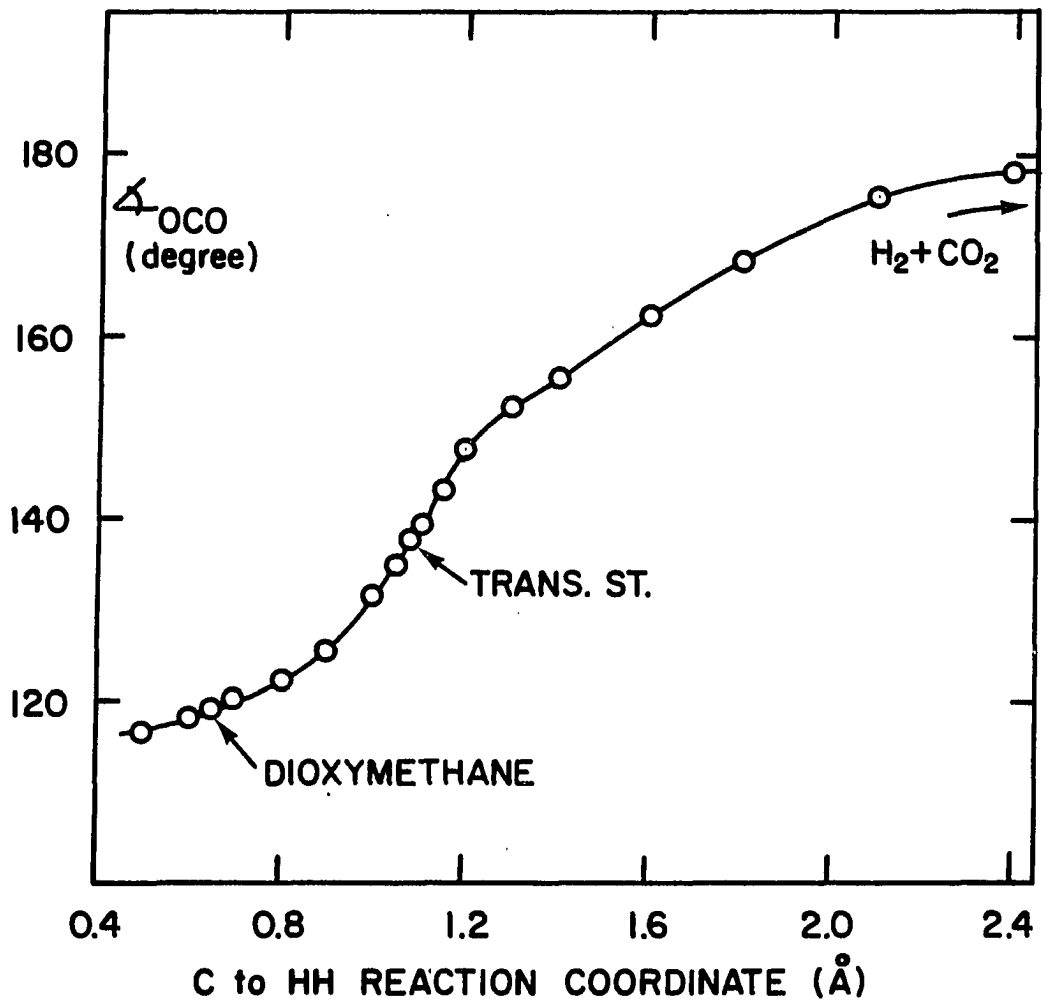


Figure 5.7. Variation of OCO angle during dioxymethane dissociation

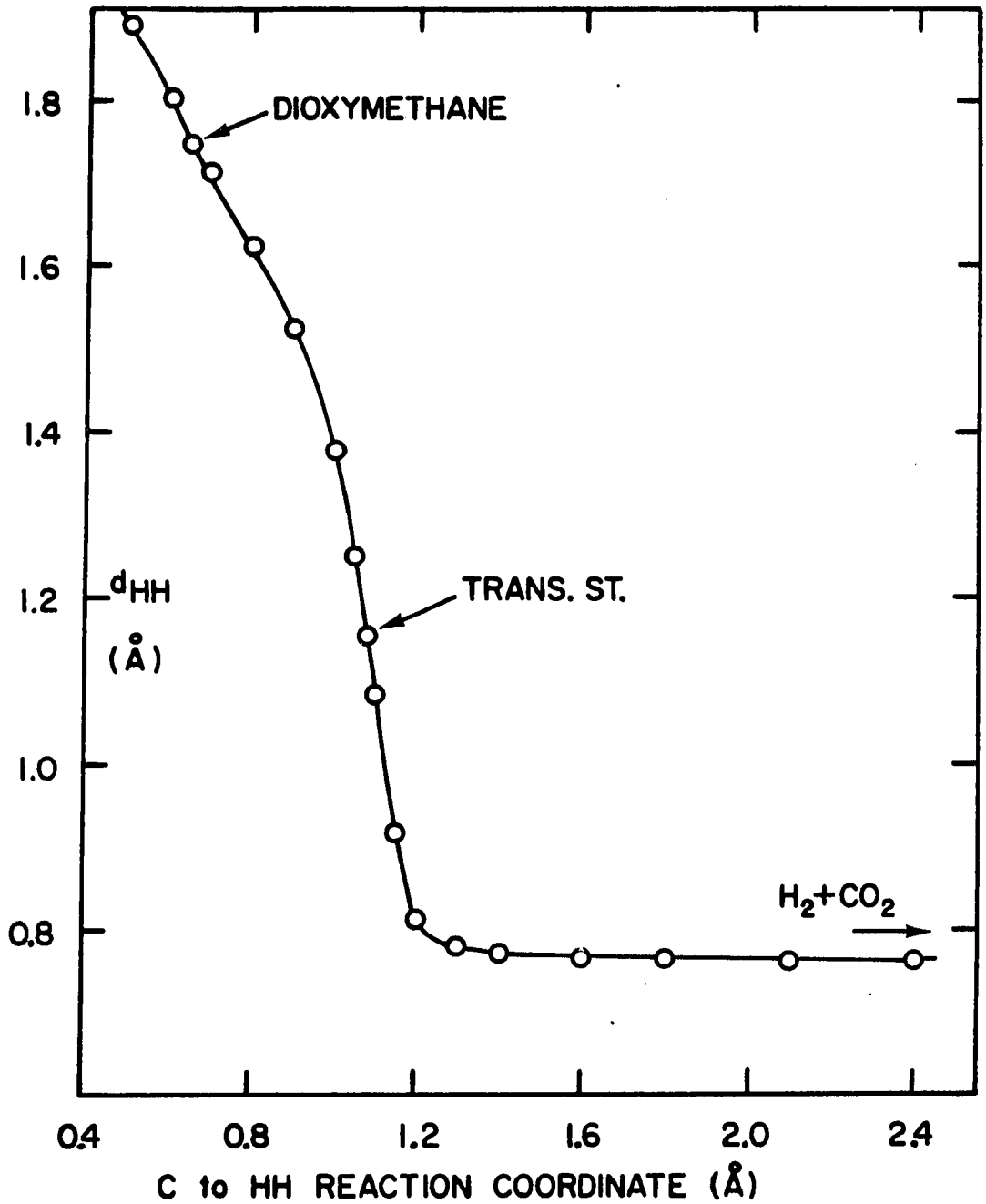


Figure 5.8. Variation of hydrogen-hydrogen distance during dioxymethane dissociation

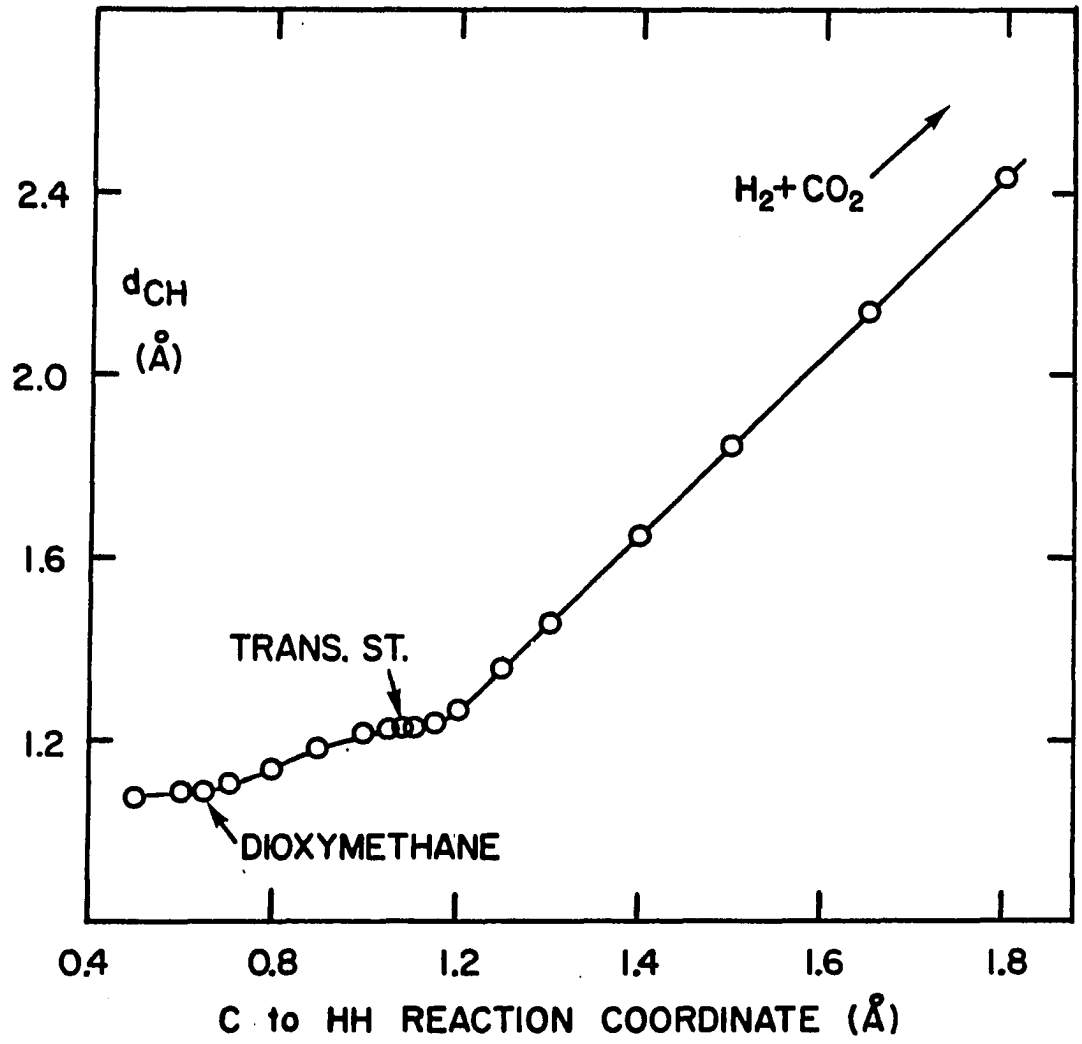


Figure 5.9. Variation of carbon-hydrogen distance during dioxymethane dissociation

Herzberg, 1979). For dioxirane, the FORS CO and OO bond lengths are 1.479 Å and 1.679 Å, respectively, compared to the experimental values of 1.388 Å and 1.516 Å (Suenram and Lovas, 1978). Thus, these bond lengths are calculated in the FORS model to be from 0.1 Å to 0.15 Å too long. These errors can be attributed to the lack of polarization functions in the basis set. Comparable errors of 0.1 Å were found by Hay and Dunning (1977) in very similar GVB-CI calculations on ozone, when polarization functions were omitted.

Dioxirane itself is well described within the SCF approximation, and the small basis used here gives an SCF geometry for dioxirane with CO and OO bond lengths of 1.434 Å and 1.518 Å, respectively. These SCF values are consistent with previous SCF calculations using small unpolarized bases (Yamaguchi et al., 1980 and Catalan et al., 1980). Similarly, the SCF bond length for singlet oxygen is 1.202 Å. The errors in the SCF bond lengths are significantly reduced from the FORS values. The elongated FORS bonds are caused by the inclusion of left-right correlation in the CO and OO sigma bonds. It should be emphasized that this correlation is absolutely essential in describing the forming and breaking of these bonds. Presumably the calculated FORS heavy atom bond lengths for the transition states to the formation of dioxirane and dioxymethane suffer the same systematic 0.1 Å or so elongation due to the lack of polarization.

For CO_2, the CO sigma bonds are not left-right correlated. These bonds are calculated to be 1.174 Å, just 0.01 Å longer than the experimental length of 1.162 Å (Herzberg, 1966). For H_2, whose bond is

correlated, the calculated length is 0.761 Å, again quite close to the experimental value of 0.741 Å (Huber and Herzberg, 1979). Presumably the geometries calculated for the dissociation reaction are more accurate than those for the methylene insertion reaction. The CO distance for dioxymethane is slightly shorter when optimized in the dissociation FRS rather than the insertion FRS.

4. Energetics

In Table 5.6 are collected the energies of all stationary points on the H_2CO_2 surface relative to the energy of the ring molecule dioxirane. In addition to the energies quoted in Tables 5.3 and 5.4, this table also includes the results for FORS calculations using the double zeta polarized basis, at the optimal geometries determined with the smaller basis. For those points which are described within the SCF approximation, SCF energy differences for both bases are included in the table, again calculated at the optimal FORS geometries. Since two different FORS functions were used, the entries for the barrier to dissociation and for $CO_2 + H_2$ are appropriately adjusted in Table 5.6. The adjustment consists of shifting the energy curve for dissociation, shown in Fig. 5.5, to bring the energy for dioxymethane into coincidence with the energy of dioxymethane in Fig. 5.1.

The FORS results using either the small or polarized basis show a small barrier (6-9 Kcal/mole) to the insertion of singlet methylene into singlet oxygen. The exothermicity to dioxirane is 105-107 Kcal/mole, the barrier to ring opening is 12 Kcal/mole, and the open form, dioxymethane, is 4-8 Kcal/mole above the ring. Note the contrast with ozone,

Table 5.6. Energetics in the C_{2v} H_2CO_2 system[a]

Molecule	SCF Results		FORS Results		KEJ[b]	CHG[c]
	Small	DZP Basis	Small	DZP Basis		
$CH_2 + O_2$	111.9	114.4	106.8	105.3		
trans. st.			113.3	114.2		
dioxirane	0.0	0.0	0.0	0.0	0.0	0.0
trans. st.			11.9	12.0	15.2	39.2
dioxymethane			4.1	8.3	1.7	25.6
trans. st.			35.5	19.9	28.3	63.0
$CO_2 + H_2$	-95.8	-114.2	-79.8	-93.0		-108.5
overall exothermicity[d]	-207.9	-228.6	-186.6	-198.4		

[a]Kcal/mole relative to dioxirane.

[b]Karlstrom, Engstrom and Jonsson (1979).

[c]Cimiraglia, Ha and Gunthard (1982).

[d]Experiment is -221.5 Kcal/mole. See text.

where the ring is the less stable form. The major change in using the polarized basis is during the dissociation of dioxymethane. The barrier to dissociation is lowered from 31 Kcal/mole to 12 Kcal/mole with the polarized basis. Finally, in the small basis, hydrogen and carbon dioxide lie 80 Kcal/mole below dioxirane, and in the polarized basis these dissociation products are 93 Kcal/mole lower than the ring.

The overall exothermicity of the reaction can be estimated. The heats of formation for methylene and carbon dioxide at zero degrees are 93.9 and -94.0 Kcal/mole, respectively (Wagman, 1968). The excitation energy to form singlet oxygen is 22.6 Kcal/mole (Huber and Herzberg, 1979). Most recent large scale CI calculations place singlet methylene 10 Kcal/mole above the triplet ground state (see, e.g., Roos and Siegbahn, 1977). The exothermicity from singlet oxygen and singlet methylene to carbon dioxide and hydrogen is, therefore,

$$\Delta H = \Delta H_f^0(CO_2) + \Delta H_f^0(H_2) - \{\Delta H_f^0(O_2, {}^3\Sigma_g^-) + T({}^1\Delta_g)\} - \{\Delta H_f^0(CH_2, {}^3B_1) + T({}^1A_1)\}$$

$$= -94.0 + 0 - (0 + 22.6) - (93.9 + 10)$$

$$= -220.5 \text{ Kcal/mole} \quad .$$

The zero point vibrational correction to this is quite small. The zero point energies are 2.1 Kcal/mole for singlet oxygen (Huber and Herzberg, 1979), 10.5 Kcal/mole for singlet methylene (Roos and Siegbahn, 1977), 7.3 Kcal/mole for carbon dioxide (Herzberg, 1966), and 6.3 Kcal/mole for hydrogen (Huber and Herzberg, 1979). The products thus have 1.0 Kcal/mole more zero point energy than the reactants, which increases the

overall exothermicity, with zero point motion excluded, to -221.5 Kcal/mole.

The calculated relative energies for the first few stationary points in Table 5.6 are quite consistent. The FORS results are changed little upon use of the polarized basis, and the relative stability of dioxirane and the reactants is not changed greatly from the SCF value. This consistency means the FORS energies from reactants up to dioxymethane are probably fairly accurate. The relative position of carbon dioxide and hydrogen is much less certain. The FORS results place the products some 20-30 Kcal/mole higher than SCF calculations. There are even big discrepancies when the same type of wavefunction is used in a larger basis. Clearly the dissociation of dioxymethane is quite exothermic, with a significant barrier of 12-31 Kcal/mole to the C_{2v} dissociation. The overall FORS exothermicity in the double zeta polarized basis is -198.4 Kcal/mole, which is rather smaller than the estimated total exothermicity of -221.5 Kcal/mole. The major source for this error is likely in the position of the final products, with the rest of the energetics shown in Table 5.5 more reliable.

The calculated barrier to C_{2v} insertion of singlet methylene is so small that a lower symmetry attack to methylene is likely to be barrierless. The remainder of the potential curve supports the Wadt-Goddard (1975) mechanism for ozonolysis. The C_{2v} ring opening of dioxirane has only a small barrier and, hence, is quite likely, as dioxirane is formed in ozonolysis with an excess of vibrational energy. The C_{2v} dissociation of dioxymethane has a somewhat larger, additional

barrier. The existence of this barrier means the breakup of dioxymethane probably proceeds through the non-C_{2v} hydrogen shift to formic acid suggested by Wadt and Goddard. This very exothermic, non-C_{2v} decay route would explain why dioxymethane has never been reported experimentally, even though it is calculated to be stable within C_{2v} symmetry.

5. Spectrum during ring opening

As dioxirane undergoes ring opening to form dioxymethane, the OO sigma bond lengthens, meaning this bond is extensively left-right correlated. Thus, the wavefunction changes from

$$\ldots \sigma^2\pi^2\pi^{*2} \qquad {}^1A_1(4\pi)$$

for dioxirane to the limiting form

$$\ldots \pi^2\pi^{*2}\{\sigma^2 - \sigma^{*2}\} \qquad {}^1A_1(4\pi)$$

just as the barrier is reached. Here σ, σ^*, π, and π^* refer to the various OO orbitals, and are, more formally, the appropriate a_1, b_2, b_1, and a_2 orbitals. For dioxymethane, the OO distance is long enough so that the σ^* orbital is nearly as low as the σ orbital. Hence, the σ orbitals can accept two electrons from the π orbitals, which suffer pair repulsions by the CH bonds. The wavefunction for dioxymethane is, therefore, dominated by

$$\ldots \sigma^2\sigma^{*2}\{\pi^2 - \pi^{*2}\} \qquad {}^1A_1(2\pi) \quad .$$

The barrier to the ring opening is caused by the avoided crossing of the

4π and 2π $^1A_1$ surfaces, and near the barrier there should be a low-lying $^1A_1$ excited state.

Near the barrier these four oxygen orbitals are quasidegenerate, and can be populated with the six electrons in a variety of ways. A large number of singlet and triplet states are possible. The most important of these is triplet dioxymethane,

$$\ldots \sigma^2\sigma^{*2}\pi^1\pi^{*1} \qquad\qquad {}^3B_2(2\pi)$$

which is the triplet state corresponding to the singlet biradical dioxymethane. Due to the small overlap of the localized orbitals containing the unpaired oxygen electrons, this triplet should be quite close to the singlet state of dioxymethane.

The triplet state corresponding to the $^1A_1(4\pi)$ state is

$$\ldots \pi^2\pi^{*2}\sigma^1\sigma^{*1} \qquad\qquad {}^3B_2(4\pi)$$

which undergoes an avoided crossing with the $^3B_2(2\pi)$ state near the barrier in the ground state singlet. Unlike the $^1A_1(4\pi)$ state, whose minimum is dioxirane, the triplet state is not expected to possess a relative minimum, as at the short bonding 00 separation of singlet dioxirane, the energy of the 00 σ^* is raised significantly.

Other low energy possibilities are

$$\ldots \sigma^2\pi^{*2}\sigma^{*1}\pi^1 + \ldots \sigma^{*2}\pi^2\sigma^1\pi^{*1} \qquad {}^{1,3}A_2(3\pi)$$

$$\ldots \sigma^2\pi^2\sigma^{*1}\pi^{*1} + \ldots \sigma^{*2}\pi^{*2}\sigma^1\pi^1 \qquad {}^{1,3}B_1(3\pi) \quad .$$

These states with three π electrons are at their lowest when all four OO orbitals are close in energy and, thus, should show minima near the barrier in the ground singlet surface. These states also favor geometries with long OO distances, keeping the σ^* antibonding orbital at low energy.

The existence of this rich spectrum was first recognized by Wadt and Goddard (1975). Vertical excitation energies for dioxymethane were reported by these workers, and also by Harding and Goddard (1978). To date, nothing concerning these states for dioxirane or at other geometries has been reported, and no experimental data are available.

Vertical excitation energies from the lowest $^1A_1$ surface at various values of the carbon to oxygen midpoint reaction coordinate to the seven excited states listed above have been calculated. The orbital space used to describe each state is the same as that used for the formation of the lowest singlet state of dioxirane and dioxymethane. That is, the four crucial oxygen orbitals are kept in the reactive set, as well as the CO bond and antibonds. The number of SAAPs in the FORS for the $^1A_2$ and $^1B_1$ states are 280, and the three triplet functions each have 384 SAAPs. The energy of each state is obtained by MCSCF optimization in the FORS space used for that state.

The low-lying singlet states are shown in Fig. 5.10, and the triplets are in Fig. 5.11. Near the barrier the eight lowest states lie within a 0.4 eV (9 Kcal/mole) range. The excited states show a discontinuity at the barrier in the lowest singlet surface. This discontinuity is due to the great difference in OO separations on either side

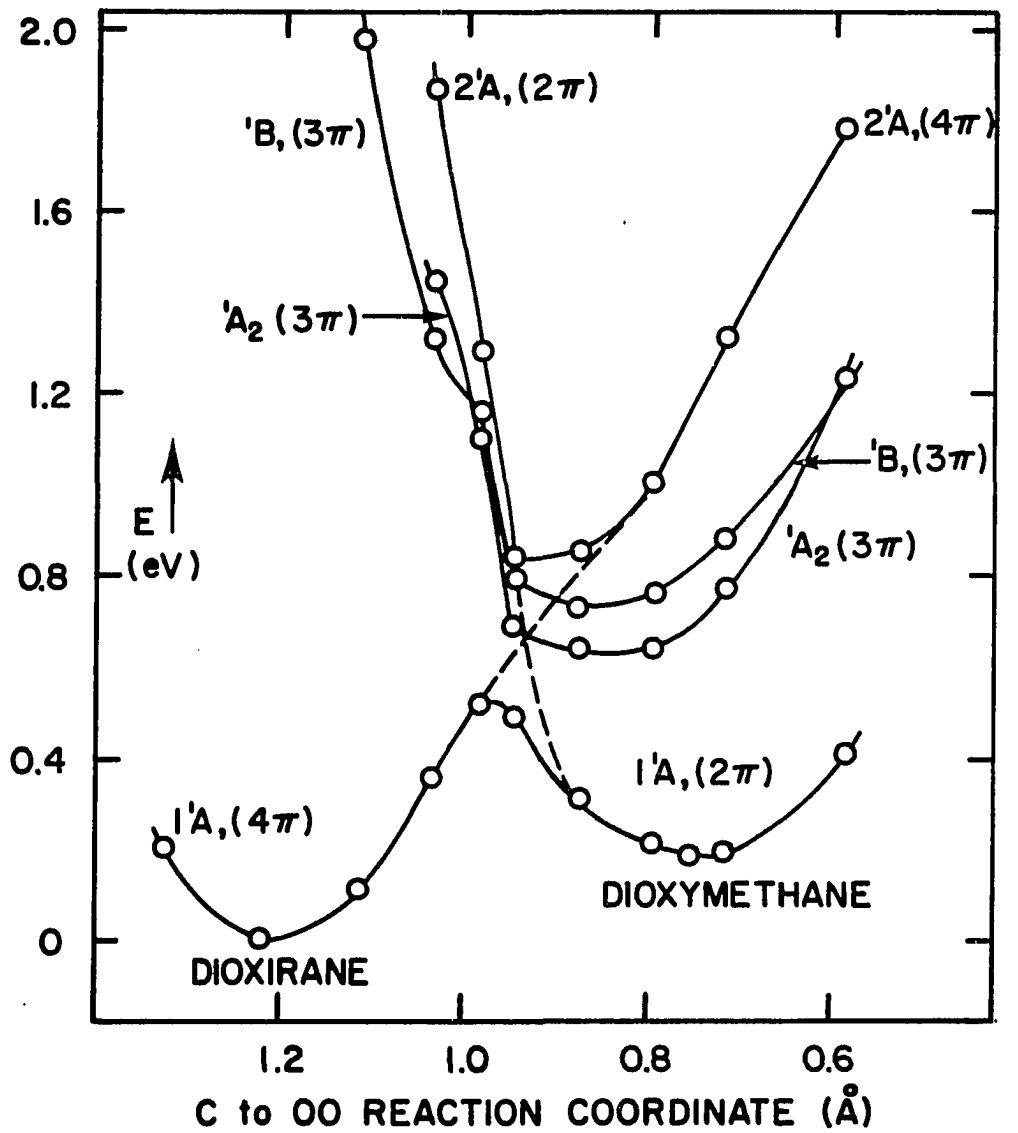


Figure 5.10. Low-lying vertical singlet states during dioxirane ring opening

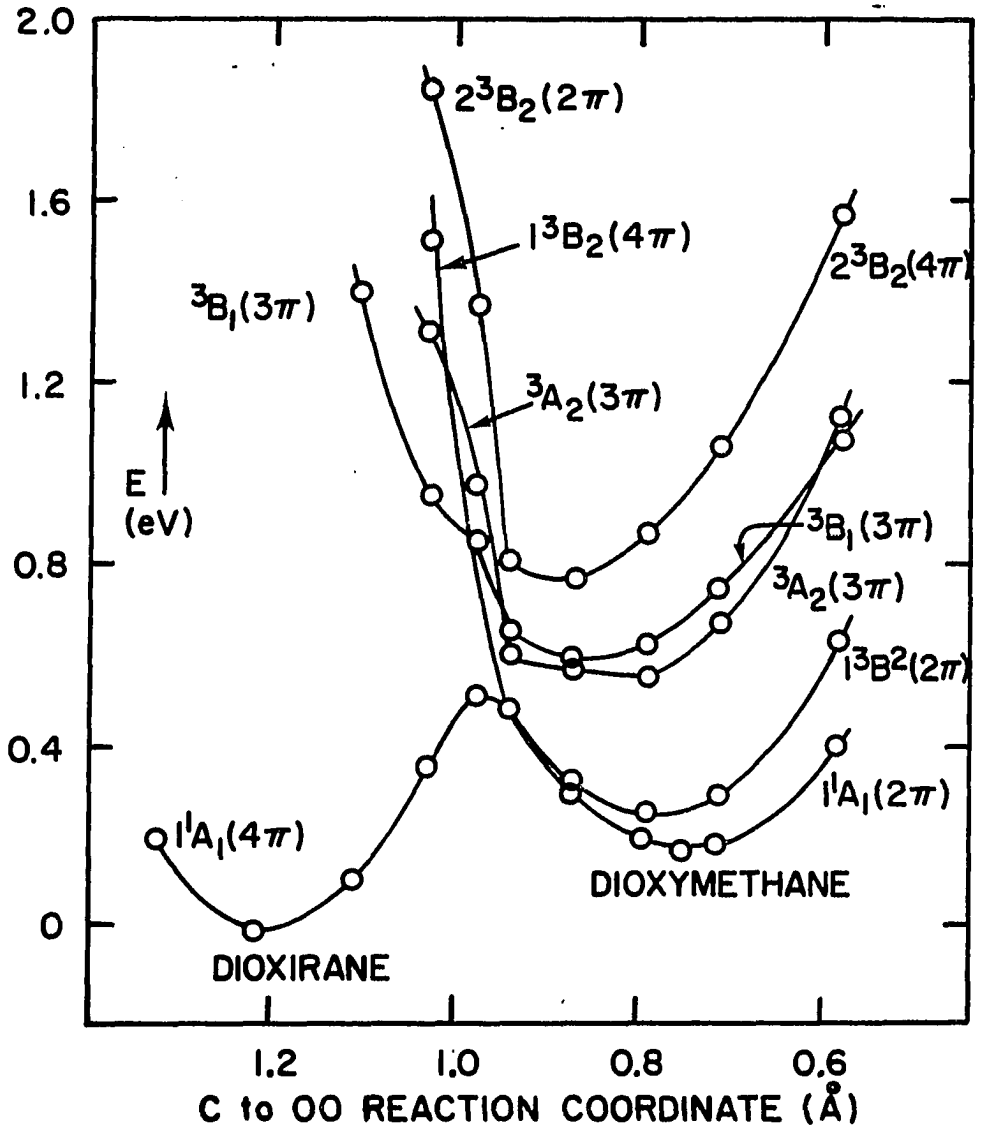


Figure 5.11. Low-lying vertical triplet states during dioxirane ring opening

of the $^1A_1$ barrier. As all of the excited states occupy the OO σ^* orbital, these states favor much longer OO distances than dioxirane has. The discontinuity is, therefore, a consequence of using the geometries of the 1^1A_1 state for all other states. The adiabatic excitation energy to these states of dioxirane should be much less than the vertical excitation energies given here.

Among all these states, the 1^3B_2 state of dioxymethane is of special interest. At the dioxymethane equilibrium geometry, its vertical excitation energy is calculated here as 2.0 Kcal/mole. The best calculation on this splitting is by Siegbahn (1979). He found, using the singlet dioxymethane geometry and the basis set of Karlstrom et al. (1979) and an MCSCF function for both singlet and triplet dioxymethane identical to that used here, that the singlet-triplet separation is 2.2 Kcal/mole. A partial first order CI calculation changed this splitting very little, to 2.0 Kcal/mole. Wadt and Goddard (1975) found this splitting to be 0.7 Kcal/mole, and Harding and Goddard (1978) report 0.9 Kcal/mole. All calculations are in agreement that the splitting is small, about two Kcal/mole at the equilibrium geometry of the singlet, and that the singlet is lowest. A FORS geometry optimization of triplet dioxymethane yields a geometry very similar to the singlet, namely a reaction coordinate of 0.766 Å, a CO bond of 1.418 Å, and a HCH angle of 111.9°. This geometry optimization lowers the triplet by just 0.5 Kcal/mole to 1.5 Kcal/mole. The triplet actually lies just below the singlet immediately after the ring opening barrier, as seen in Fig. 5.11.

Near the ground state barrier the seven excited states given here are all quite close to the ground singlet surface, and within 0.9 eV of the dioxirane energy. Not all possible occupations of the four OO orbitals with the six electrons are included here. These additional states have been examined as higher roots of the lower states, and some MCSCF optimizations on these higher states. There is a large gap between the states shown in Figs. 5.10 and 5.11 before these additional states are reached. Near the barrier these additional states form a group lying some 4-5 eV above the dioxirane energy, and higher elsewhere. Some of these states involve excitations from the CO bonds.

6. Comparison to previous work

In a previous theoretical study of the conversion of dioxirane to dioxymethane, Karlstrom et al. (1979) used an MCSCF wavefunction consisting of the same ten configurations, all double replacements, at all points on the potential surface between dioxirane and dioxymethane. These ten configurations are all contained in the 320 SAAP reaction space used here to describe the insertion reaction and formation of dioxirane. In view of the large changes in the dominant portion of the FORS wavefunction between dioxirane and dioxymethane, this means Karlstrom et al. (1979) represented dioxirane and dioxymethane with only five configurations each. This is an inadequate number of configurations to represent these molecules, but the limitation to double replacements mean that these are not even the most important five configurations for dioxirane or dioxymethane. The ten configurations chosen are even less adequate near the conversion barrier, where configurations with four open

shell electrons are important. In the approach used here, (i) the MCSCF configurations are found in the full reaction space, rather than selected according to preconceived notions, and (ii) the energy is calculated to the accuracy of the full reaction space, rather than a small MCSCF function. Such an approach is clearly more satisfactory.

Karlstrom et al. (1979) do include the four most important configurations, given in the preceding section, needed for the ground state totally symmetric singlet in their ten configuration MCSCF function. Consequently, their energy results, as seen in Table 5.6, are comparable to those obtained here. The major difference is that calculation in the 320 configuration full reaction space used here places dioxymethane 8 rather than barely 2 Kcal/mole above the ring. Karlstrom et al. (1979) used a double zeta plus polarization basis and, consequently, obtained a more accurate geometry for dioxirane than the FORS structure reported here.

Karlstrom et al. (1979) also extend their examination of the potential surface with their ten configuration MCSCF function up to the transition state for the dissociation of dioxymethane to CO_2 and H_2. None of the configurations used allows for the increased left-right correlation in the CH bonds as they elongate during this dissociation, or in the incipient HH bond, as is the case for the 320 configuration full reaction space used to describe this dissociation in the present work. However, the smooth transformation from dioxymethane to its dissociation products means that Karlstrom et al. (1979) do catch the

essence of the dissociation, which is the change from a two to a single dominant configuration.

Cimiraglia et al. (1982) have also considered the ring opening to dioxirane to form dioxymethane, as well as the dissociation of dioxymethane to carbon dioxide and hydrogen. This study uses orbitals obtained from SCF calculations, followed by a multireference CI calculation to recover correlation effects. In view of the strong configuration mixing near the conversion barrier, and the the two configuration, singlet biradical nature of dioxirane, the restriction of the orbital optimization to an SCF function by Cimiraglia et al. (1982) is inappropriate. As seen in Table 5.6, this causes their energies for the barrier to ring opening and for dioxymethane to be too high.

Cimiraglia et al. (1982) perform only a limited geometry variation during the dissociation of dioxymethane. They linearly interpolate the geometry between dioxymethane and its dissociation products, assuming the reaction is complete when CO_2 and H_2 are four angstroms apart. As shown here, the dissociation is complete when these molecules separate by 2.4 Å. The poor geometries chosen by Cimiraglia et al. (1982) lead to a very large barrier for the dissociation reaction. The exothermicity between dioxirane and CO_2 plus H_2 reported by these authors may well be correct. Both of these structures are well-described by the SCF plus multireference CI wavefunction used by these authors, and as noted above the FORS calculations here seem to place H_2 and CO_2 too high.

E. Conclusions

The entire least motion C_{2v} potential surface for the H_2CO_2 system has been investigated. The FORS calculations show that the formation of dioxirane from singlet methylene and singlet oxygen, as might be expected, is very exothermic, by 105 Kcal/mole. The barrier to the insertion of methylene is quite small, 9 Kcal/mole, so that a lower symmetry attack is likely barrierless.

The ring opening of dioxirane to form dioxymethane is crucial to the Wadt-Goddard (1975) mechanism for gas phase ozonolysis. The calculations here reveal the ring opening is slightly endothermic, by 4-8 Kcal/mole, with a barrier only slightly higher, 12 Kcal/mole. This contrasts with the isoelectronic system ozone, where the open form lies 28 Kcal/mole below the ring form (Hay and Dunning, 1977). Nonetheless, the ring opening postulated for dioxirane is certainly feasible. The vertical excitation energies near the ring opening barrier are reported for all states within 4 eV of dioxirane.

The final reaction considered is the dissociation of dioxymethane to molecular hydrogen and carbon dioxide, one of the observed products of gas phase ozonolysis of ethylene. The barrier to the C_{2v} dissociation of dioxymethane is 12-31 Kcal/mole, so that the presumably barrierless hydrogen shift in dioxymethane to formic acid, as suggested by Wadt and Goddard (1975), is the more likely decomposition route. As expected, carbon dioxide and hydrogen are much more stable than dioxirane, by 80-93 Kcal/mole. The calculated energetics for this dissociation are probably less accurate than those for the insertion and ring opening.

The investigation reported here features more extensive geometry searching than previous work. Moreover, the configurations are selected by the FORS model appropriate to the portion of the surface being investigated. The configurations chosen to treat this system are, thus, more numerous and less arbitrarily chosen than previous investigations of parts of this potential surface (Karlstrom et al., 1979 and Cimiraglia et al., 1982). The calculations reported here support the decomposition mechanism for dioxirane in gas phase ozonolysis suggested by Wadt and Goddard (1975).

VI. THE INTRAATOMIC CORRELATION CORRECTION TO THE FORS MODEL: APPLICATION TO H_2, NH, AND F_2

A. Theoretical Development

Chemists have long thought of molecules as composed of atoms situated in close proximity to each other. This view was adopted during the early stages of molecular quantum mechanics, for instance, in valence bond theory or in the LCAO approach to constructing MOs. As the minimal basis of Slater functions gave way to the modern practice of using extended Gaussian bases, contracted in segmented fashion, and as molecular orbital theory supplanted valence bond theory, the participation of atoms in a molecular wavefunction has become blurred. As shown in Chapter II, the FORS model restores the concept of the atomic minimal basis in the framework of extended basis calculations. Appropriate manipulations of FORS wavefunctions reveal the manner in which atoms participate in molecular binding. This atomic analysis of FORS wavefunctions permits an empirical correction to energies calculated by the FORS model.

1. Synopsis of AIM theory

Moffitt (1951b) recognized that the errors in the first, rather primitive molecular calculations, while only one percent or so of the total energy, were still larger than most important chemical energy differences such as bond energies and excitation energies. Moffitt realized that it was nearly impossible to improve the computations of that period to remove these errors, and so proposed an empirical

correction scheme, known as the Atoms in Molecules (AIM) method. The scheme involves partitioning the molecular Hamiltonian into interatomic and intraatomic terms. The intraatomic terms are much larger, and are known from spectroscopic data to greater accuracy than they can be calculated. The interatomic terms are calculated for the molecule of interest, and because they are smaller than the intraatomic terms, are presumably calculated to greater accuracy. Moffitt (1951a) was able to apply this synthesis of experimental and calculational data to obtain potential curves of improved accuracy for several valence states of molecular oxygen. Further exploration of the concept of atomic valence states in molecules is given by Moffitt (1954).

The AIM method will be sketched for diatomic molecules AB. The molecular wavefunction is built from valence bond structures for the diatom, such as AB, A^*B, AB^*, A^+B^-, A^-B^+, etc. To avoid repetition, the ground and excited states of the neutral atoms, ions, or multiply charged ions, will be collectively referred to as atomic states. A convenient notation for a given valence bond structure, termed a composite function (CF) by Moffitt (1951b) is

$$|A_iB_j\rangle = \tilde{A}|A_i\rangle|B_j\rangle \quad , \tag{6.1}$$

where $\tilde{A}$ is the coset antisymmetrizer that produces a totally antisymmetric function from the product of antisymmetric atomic functions $|A_i\rangle$ and $|B_j\rangle$. Of course, for a neutral molecule, the atomic states must have equal but opposite charges. A minimal basis is used to form the various atomic states. As the orbitals on each atom are nonorthogonal,

the calculation of such a valence bond function involves the difficult computation of Hamiltonian and overlap integrals between the CFs:

$$\begin{aligned} H_{ij,k\ell} &= \langle A_i B_j | H | A_k B_\ell \rangle \\ S_{ij,k\ell} &= \langle A_i B_j | A_k B_\ell \rangle \quad . \end{aligned} \tag{6.2}$$

The difficulty of calculating such matrix elements in a nonorthogonal basis has caused the virtual abandonment of the valence bond theory in favor of molecular orbital theory. However, the clear identity of atomic states in valence bond wavefunctions facilitates the use of the AIM correction method.

The derivation of the AIM correction procedure is based on partitioning the molecular Hamiltonian into intra- and interatomic terms

$$H = H^A + H^B + V^{AB} \quad . \tag{6.3}$$

The intraatomic term for atom A is

$$H^A = -\frac{1}{2}\sum_i^A \nabla_i^2 - \sum_i^A \frac{Z_A}{R_{Ai}} + \frac{1}{2}\sum_i^A \sum_j^A \frac{1}{r_{ij}} \quad , \tag{6.4}$$

where the sums run only over the electrons possessed by atom A (this number changes if in a particular CF the "atom" is an ion) and H^B is defined similarly. The interatomic interaction operator is

$$V^{AB} = -\sum_i^B \frac{Z_A}{R_{Ai}} - \sum_i^A \frac{A_B}{R_{Bi}} + \sum_i^A \sum_j^B \frac{1}{r_{ij}} + \frac{1}{R_{AB}} \quad . \tag{6.5}$$

The Hamiltonian matrix element between two CFs composed of <u>exact</u> atomic eigenfunctions is

$$
\begin{aligned}
H_{ij,k\ell} &= \langle A_i B_j | H^A + H^B + V^{AB} | A_k B_\ell \rangle \\
&= (E^A_k + E^B_\ell) S_{ij,k\ell} + V^{AB}_{ij,k\ell} \quad .
\end{aligned}
\tag{6.6}
$$

The Hermitian conjugate element is

$$
\begin{aligned}
H_{k\ell,ij} &= \langle A_k B_\ell | H^A + H^B + V^{AB} | A_i B_j \rangle \\
&= (E^A_i + E^B_j) S_{ij,k\ell} + V^{AB}_{k\ell,ij} \quad .
\end{aligned}
\tag{6.7}
$$

These expressions are equal, as the total Hamiltonian is self-adjoint. The metric S is also Hermitian, while the interaction operator itself is not. Equations (6.6) and (6.7) are derived without approximation for exact atomic eigenfunctions.

The essence of the atoms in molecules method is that the interatomic term V^{AB} and the overlaps S can be calculated much more accurately than the much larger atomic energies E. Approximating the atomic states by expanding in some appropriate basis, and denoting quantities calculated in this basis by a tilde, the two matrix elements above are approximately given by

$$
\begin{aligned}
H_{ij,k\ell} &\simeq (E^A_k + E^B_\ell) \tilde{S}_{ij,k\ell} + \tilde{V}^{AB}_{ij,k\ell} \\
H_{k\ell,ij} &\simeq (E^A_i + E^B_j) \tilde{S}_{ij,k\ell} + \tilde{V}^{AB}_{k\ell,ij} \quad .
\end{aligned}
\tag{6.8}
$$

These are no longer equal, due to the approximate nature of the atomic state functions, and a Hermitian matrix is obtained by averaging these conjugate elements,

$$H_{ij,k\ell} \simeq \frac{1}{2}(E_i^A + E_j^B + E_k^A + E_\ell^B)\tilde{S}_{ij,k\ell} + \frac{1}{2}(\tilde{V}^{AB}_{ij,k\ell} + \tilde{V}^{AB}_{k\ell,ij}) \quad . \quad (6.9)$$

Calculation of the matrix elements of the interaction operator V^{AB} defined in Eq. (6.5) is exceedingly difficult. Following Moffitt's (1951b) second alternative, these elements are obtained approximately by calculating all quantities in Eq. (6.6) in the basis representation,

$$\tilde{V}^{AB}_{ij,k\ell} \simeq \tilde{H}_{ij,k\ell} - (\tilde{E}_k^A - \tilde{E}_\ell^B)\tilde{S}_{ij,k\ell} \quad , \quad (6.10)$$

and similarly for $\tilde{V}^{AB}_{k\ell,ij}$. Substituting into (6.9) gives

$$H_{ij,k\ell} \simeq \tilde{H}_{ij,k\ell} + \frac{1}{2}\{(E_i^A - \tilde{E}_i^A) + (E_j^B - \tilde{E}_j^B) + (E_k^A - \tilde{E}_k^A) + (E_\ell^B - \tilde{E}_\ell^B)\} \times \tilde{S}_{ij,k\ell} \quad . \quad (6.11)$$

This is the Atoms in Molecules (AIM) Hamiltonian and is trivially calculated provided that the usual $\tilde{H}$ and $\tilde{S}$ matrix elements needed for a conventional calculation in the CF basis are known. The second term in the AIM Hamiltonian represents the corrections for the errors involved in calculating the large intraatomic energies. The partitioning of the Hamiltonian in Eq. (6.3) is purely formal, to facilitate the development of the approximation (6.11). It is not necessary to calculate matrix elements of the interaction operator V^{AB}, as Moffitt (1951b) suggests as his first alternative.

2. Types of atomic correction terms

The essence of any AIM method, and the differences between them, are the type corrections used in applying (6.11). In Moffitt's (1951a) original application, AIM calculations on oxygen are performed in a

minimal basis of Slater orbitals, with screening constants selected by Slater's rule. The difference between the exact energy E of an atomic state, and the SCF energy calculated in such a basis, can be attributed to several causes:

i) failure to optimize the screening parameter to achieve the lowest possible atomic energy,

ii) failure to expand the atomic minimal basis set in terms of several Slater, or Gaussian functions, i.e., to use "Hartree-Fock" AOs,

iii) additional errors in the excited states or ions due to using screening constants, or contractions appropriate to the neutral atom ground state,

iv) atomic correlation effects discussed further below, and

v) omission of relativistic effects from the Hamiltonian.

An additional source of error in molecular calculations is the failure to use an extended basis set in the molecular calculation, that is a more flexible contraction than the minimal basis, as well as polarization functions.

In the sequel, a variant of the AIM method will be applied to F_2. To illustrate the sizes of the correction terms listed above, consider the ^2P ground state of the fluorine atom. The energy of F in a minimal basis, with exponent chosen by Slater's rules, is -98.9317 h. Optimization of the exponent lowers this by 0.0104 h. More dramatic lowering is produced by increasing the basis to the atomic limit; this lowers the energy by 0.5328 h. Atomic correlation for F lowers the energy by

0.328 h, and relativistic terms produce a lowering by 0.0805 h. The corrections sum to 0.9517 h, and represent 0.95% of the total energy of F, -99.8834 h. To put the size of these corrections in perspective, the bond strength of F_2 is just 0.061 h. The third type of error above depends on the particular state considered. The error in using near Hartree-Fock F atom ground state orbitals (14s,7p/2s,1p) for the ^2S excited state of F is 0.0190 h, for ^1S F^-, 0.1518 h, and for ^3P F^+, 0.1075 h.

All AIM calculations to date have been performed with minimal primitive sets, or minimal contractions, rather than extended bases. The reason for this is that the MBS, containing just one function per AO, permits ready identification of orbital occupancies and, hence, atomic states in the molecular wavefunction. An s^2p^N occupation is easily distinguished from a s^1p^{N+1} configuration, etc. The final type of error mentioned above, namely using a minimal contraction of Hartree-Fock quality versus an extended basis for F_2, divided by two for comparison to the atomic terms, is 0.019 h. This error can be substantially larger for other first row diatomics, and unlike the other errors is purely molecular. The extended basis contributes directly to $\tilde{H}$ in Eq. (6.11), rather than affecting the correction term in that equation. No previous AIM calculations have been able to account for this type of basis truncation error.

If the five purely atomic terms are applied in Eq. (6.11), the correct positions of all valence molecular eigenstates are obtained in the limit of infinite separation of the atoms. In this limit, the overlap

matrix becomes a unit matrix, and the correction terms by construction, adjust the Hamiltonian $\tilde{H}$ so as to asymptotically yield the correct state energies. It is to be hoped that at finite bond lengths, the potential curves given by the corrected Hamiltonian have the correct shapes and remain correctly spaced with respect to each other. The success of the AIM method must be measured by the extent to which this is true.

Further clarification of the role of atomic correlation in the AIM method is needed. The composite functions are usually chosen here, and for the most part in other AIM applications, as products of single configuration atomic valence states. Direct computational recovery of the correlation error, say for the F atom, would involve excitation into the 3s, 3p, 3d, etc. higher atomic orbitals. This atomic correlation is, therefore, entirely "dynamical" in nature, in that its recovery in a CI calculation would require excitations into the external orbital space. This atomic correlation is not the only type present in the molecule, however. In keeping with the spirit of the FORS model in the present work, all possible CFs with the appropriate overall symmetry which can be formed by taking products of all atomic (including ionic) valence states will be chosen as the base function to which the AIM correction will be applied. In much of the previous AIM work, only the more important CFs were used. The CFs used here differ only in the way the valence, or internal orbitals, are occupied, and so the molecular correlation energy recovered by the uncorrected function is of the "near degeneracy" type. In light of this, Eq (6.11) can be interpreted in the following way: the base function itself recovers all the "near

degeneracy" correlation through the uncorrected Hamiltonian, and the correction term recovers the "dynamical" molecular correlation, which is interpreted in the AIM scheme as purely atomic in nature.

Two points in the preceding paragraph deserve further discussion. The uncorrected wavefunction is defined above to consist of all possible CFs formed from all atomic states. In fact, if the atomic basis in which this calculation is performed is sufficiently flexible, and the AOs used to construct the CFs are optimized in that basis, this definition is equivalent to the FORS wavefunction defined from the molecular orbital viewpoint in Chapter II. A minor point involves the atomic states whose products are the CFs. In most instances, these atomic states are of the SCF type, that is, they consist of a single configuration. In a few instances, a multiconfiguration function must be used, e.g., when the two configurations s^2p^N and s^0p^{N+2}, ($n \leq 4$) happen to possess the same state symmetry. This choice is consistent with the above discussion. The atomic multiconfiguration function recovers the "near degeneracy" correlation of the atom, and the remaining correlation error is "dynamical" in nature and is treated as a correction.

The types of atomic correction energies given above fall into two general categories: basis set errors, and correlation or relativistic effects. With the aid of the new digital computers, the focus of quantum chemistry from the late 1950s onward was on the elimination of these basis and correlation errors. With modern computers and programs, basis errors can be reduced practically to zero for diatomics and are quite small even for moderately sized molecules. Recovery of the

correlation error has proved the more difficult problem, with most MCSCF or CI procedures aimed primarily at the recovery of valence shell correlation errors. The FORS calculations on diatomic molecules presented at the end of Chapter II serve as a perfect example of these decades of computational improvement: the MCSCF procedure recovers all "near degeneracy" type correlation, and the estimated basis truncation errors are smaller than the errors in the calculated bond energies. Unfortunately, these bond energy errors are several times larger than chemical accuracy, and so one must look primarily to the atomic correlation, and possibly to the neglect of relativistic terms in the Hamiltonian to obtain more accurate bond energies.

3. The ICC modification of AIM theory

Shortly after Moffitt's original AIM proposal, a number of workers were attracted to the concept by its promise of greatly improved accuracy for little computational effort. Only one of the resulting AIM variants will be discussed here. Review articles, with reference to other work, have been given by Arai (1960), Hurley (1963), and Balint-Kurti and Karplus (1974). The development of the AIM approximation in Eq. (6.11), and the notation used here, follows closely the review by Balint-Kurti and Karplus (1974), who obtain (6.11) with more rigorous attention to the commutation properties of the interaction operator and the antisymmetrizer, and with more discussion of the approximations involved in its development. Verhaegen and co-workers have developed a correlation correction scheme based on Mulliken population analyses of MBS SCF functions that can distinguish only the weights of atomic

configurations (not individual atomic states) in the wavefunction. This scheme is similar in spirit to AIM methods but is quite different in its application, and is most recently described by Lieven, Breulet and Verhaegen (1981).

Application of Moffitt's original AIM proposal was shown by several workers to yield binding energies larger than experiment, that is, the original underestimates of these bond energies were being overcorrected by the AIM correction. Hurley (1956) attributed these overcorrections to the large corrections applied to those CFs formed from ions or excited states built from orbitals with exponents chosen for the ground state atoms. In the most successful modification of AIM theory to date, Hurley (1956) proposed the Intraatomic Correlation Correction (ICC) scheme. The ICC correction is obtained from atomic or ionic calculations in which the exponents are reoptimized for each such state, that is, corrections of type iii above are ignored. The resulting correlation correction for such a state is applied unchanged in the molecule, even though all CFs are constructed using the same screening parameters. An interesting and highly useful consequence of the ICC correction is that the screening parameters can be taken as those giving the lowest molecular energy rather than those best for the ground state neutral atoms, since the ICC correction is assumed independent of the screening constants used in the molecular calculation. As most of Hurley's work was performed with minimal Slater bases, these excited states and ions as well as the neutral atoms still possess a large correction due to failure to expand these states in a larger primitive set, the type ii

error given above. In later work, Hurley (1963) uses Hartree-Fock AOs rescaled for the molecule to construct CFs, thereby substantially reducing these type ii contributions to the ICC corrections.

The assumption that the correlation correction is independent of the screening parameters (or the scaling of a contraction) at first glance seems quite arbitrary. However, a number of workers have shown that correlation effects are surprisingly independent of orbital size and shape. Hurley (1956) gives this physical motivation for the assumption: the correlation energies of H^-, He, and Li^+ (including also the MBS truncation error) are 1.49, 1.51, and 1.54, respectively, although the Slater MBS SCF functions for these three species have quite different screening constants. The pure correlation errors for two electron atoms are likewise almost constant, changing from 1.08 eV for H^-, 1.14 eV for He, 1.18 eV for Li^+, to 1.24 eV for Ne^{+8}. Miller and Ruedenberg (1968) have shown the 1s inner shell correlation for the four electron atoms is 1.12 eV for Li^-, 1.07 eV for Be, and 1.00 eV for Ne^{+6}. As another example, take the isoelectronic series of nine electron atoms: the total correlation energy of F is 8.82 eV; Ne^+, 8.92 eV; and Na^{+2}, 9.14 eV (Verhaegen and Moser, 1970). The correlation error is remarkably constant for this sequence, especially in view of the fact that the SCF orbitals deform (shrink) more in proceeding from F to Na^{+2} than they do in going from F to excited F, F^+, or F^-. It, thus, seems physically justified to take the correlation energies as independent of orbital sizes. Of course, correlation energies do vary with the number

of electrons occupying these orbitals, or the manner in which these electrons are coupled.

One drawback of the ICC procedure is that unlike the AIM scheme, it does not correctly predict the positions of all valence states of a diatom in the limit of large bond lengths. The neglect of the type iii correction term improves the shape of the potential curve for which the screening parameters and atomic contractions were adjusted, at the expense of the correct asymptotic behavior for the other states. The curves for other states must be obtained by separate ICC calculations on each desired state, rather than obtaining all states from the same secular equation, as in the original AIM formalism.

4. The FORS-IACC procedure

As shown in Chapter II, the FORS bond energies for diatomic molecules possess an error rather larger than chemical accuracy. The Atoms in Molecules concept with a correction scheme similar to the ICC scheme can be applied to the FORS wavefunction to remove a goodly portion of this error. The correction scheme used here is termed the IntraAtomic Correlation Correction (IACC) to show its commonality with the ICC approach, but differs from it in two important fashions.

The major innovation in the FORS-IACC scheme is that the CFs are constructed from orthonormal "atomic" orbitals. As shown in Chapter III, the MOs of a FORS wavefunction may be localized according to a projection procedure in such a fashion that the localized orbitals closely resemble unhybridized atomic orbitals. When constructing CFs, the localized MO most closely resembling a particular AO is used in that AO's

stead. Examination of the projected localized MOs in Figs. 3.1, 3.3, and 3.6 shows the extent to which this replacement is justified. (Note that, in general, the projected MOs used as the three p, or five d, etc. orbitals are not spatially equivalent.) The use of these projected localized MOs in the CFs permits a great simplification of the AIM approximation. Since the orbitals are orthonormal, the CFs are also orthonormal at <u>all</u> bond distances. Equation (6.11) simplifies to

$$H_{ij,k\ell} \cong \tilde{H}_{ij,k\ell} + \{(E_i^A - \tilde{E}_i^A) + (E_j^B - \tilde{E}_j^B)\}S_{ij,k\ell} . \quad (6.12)$$

The computation of $\tilde{H}$ is also considerably simplified by orbital orthonormality. The correction term may be viewed as a perturbation which is diagonal in the CF representation.

A second important feature of the FORS-IACC procedure is that the uncorrected Hamiltonian is computed using an extended, flexibly contracted atomic basis set. The minimal basis orbitals to which electrons are assigned in forming CFs are found by localization of the FORS MOs, after these have been optimally determined in the QBO basis. The three types of atomic basis errors and the purely molecular error of inflexible contraction described in Section VI.A.2 are minimized by taking the molecule optimized PLOs as the atomic minimal basis set for the construction of CFs. Because the FORS calculations, at least for diatomic molecules, are performed in large QBO sets, the errors due to finite basis expansion are small and need not be accounted for in the correction term in Eq. (6.12). Therefore, the IACC correction term used in applying (6.12) consists only of correlation and relativistic corrections.

The correlation correction used in the IACC scheme is a generalization of the ICC correction, for which correlation errors are assumed independent of orbital scaling. Here the orbitals are determined by molecular optimization in a flexible basis and, hence, are deformed from the neutral free atom SCF AOs in a more complicated fashion than just rescaling. However, the PLOs were shown in Chapter III to be quite similar to the free atom SCF AOs, with overlaps usually in excess of 0.9. Therefore, the IACC correlation and relativistic correction for an atomic state built from PLOs are assumed to be equal to that of the same state formed from its optimal free atom AOs.

Given the Hamiltonian matrix in the CF representation, application of the IACC correction requires only knowledge of the correlation and relativistic energy corrections to the various atomic states occurring in the CFs. Happily, there is much more information of this type available now than in the early days of AIM theory. There are two key compilations of these data for the three to ten electron atoms (Li to Ne), and for many isoelectronic positive ions as well. Correlation energies, in some cases corrected for the 2s-2p near degeneracy, for all states arising from the $2s^2 2p^N$ configurations, are given by Verhaegen and Moser (1970). Correlation energies for all $2s^1 2p^N$ and $2s^0 2p^N$ states, and relativistic corrections for all three types of configurations are given by Desclaux, Moser and Verhaegen (1971). For higher elements, less data are available. Clementi (1965) has given correlation corrections for the ground states of the first 22 elements, and some of their positive ions. Fraga et al. (1976) give relativistic

corrections to Hartree-Fock energies for the ground states of the first 102 elements and many of their positive ions

The availability of correlation and relativistic corrections for negative ions is considerably less, owing largely to the difficulty in obtaining experimental data for these species. Generally only the singly negative ion, at most, will be stable. Electron affinities are difficult to measure accurately, and are normally only available for the ground states of the ions. Hotop and Lineberger (1975) have given experimental electron affinities for the first 85 elements of the periodic table, including a very few negative atom excited states. Schaefer et al. (1969) have given less accurate theoretically derived electron affinities for the atoms B to F, including a number of the low-lying excited negative ion states. Correlation and relativistic corrections can be extracted from these electron affinities, with the help of judicious extrapolation of the known corrections for the neutral and positive ions. Clementi and Roetti (1974) give SCF energies for singly negative ions from Li^- to I^-.

Only singly negative ions occur in the diatomic molecules considered in the present case. However, highly negative ions can occur, e.g., $C^{-4}O^{+4}$ CFs would be found included in the CO wavefunction. For doubly negative ions, there are virtually no experimental data (Hotop and Lineberger, 1975), and extrapolation to these or even more negative ions from the positive and neutral species is dubious. It is recommended that the correction applied to any such CF for which accurate values

are unknown be taken as equal to the correction applied to the CF formed with neutral atoms in their ground states.

As a result of the resolution of the errors of atomic Hartree-Fock energies into correlation and relativistic contributions by Desclaux, Moser and Verhaegen (1971), it is possible to attribute the molecular errors to either correlation or relativistic contributions. While theoretical chemists usually believe relativistic effects to be much more constant than correlation energies, relativistic corrections to atomic energies are hardly negligible. Recently, Feller and Davidson (1980) have suggested the discrepancy between experimental and large scale CI determinations of the singlet-triplet separation in methylene is due to relativistic contributions to the carbon atom, said to be in the s^2p^2 or sp^3 valence state, respectively. The IACC scheme permits the relative importance of correlation versus relativistic errors in molecular calculations to be estimated.

B. Applications

1. Implementing the IACC correction

To apply the IACC correction embodied in Eq. (6.12) to a FORS wavefunction, the Hamiltonian matrix elements between CFs formed from PLOs are needed. Although these matrix elements could be evaluated directly, two difficulties arise. The specification of just which CFs occur in the full valence space is not as simple as specifying which SAAPs occur in the FRS. Furthermore, while existing ALIS programs can readily evaluate Hamiltonian matrix elements between SAAPs, this is not

the case for the calculation of these matrix elements for CFs. While both these problems are soluble, in the present case existing programs are used to do as much computation as possible in a SAAP basis before transforming to a CF representation.

As already discussed, the full set of CFs constructed from all atomic valence states is equivalent to the full valence space of SAAPs described in Chapter II. These two bases are related by an orthogonal transformation,

$$\{CFs\} = \{SAAPs\}T \quad . \tag{6.13}$$

The full valence space in the SAAP basis is readily generated (Lam, 1982), as is the Hamiltonian in this SAAP basis. The Hamiltonian in the CF basis is obtained by similarity transformation,

$$H^{CF} = T^{\dagger} H^{SAAP} T \quad . \tag{6.14}$$

The transformation T in Eq. (6.13) is a sparse matrix, and its elements are simply defined numbers. T is independent of the molecular geometry, and its computation is illustrated by an example below. The algebraic manipulations needed to obtain this transformation are the only difficult step in applying the IACC procedure. A program to carry out this transformation, and to diagonalize the CF Hamiltonian before and after applying the IACC correction has been written. This program also calculates the correction to the molecular energy via first order perturbation theory.

In transforming SAAPs built from PLOs to CFs constructed from the same PLOs, it is necessary to know how the PLOs are occupied in the

various atomic states. The atomic states must be expressed in terms of real rather than complex atomic orbitals, as the PLOs are of course real. All valence states arising from the p^N configurations, $1 \leq N \leq 6$, have been found for real atomic p orbitals. These states, and the method by which they were obtained, are given in the Appendix to this chapter.

A difficulty faced in obtaining the transformation to the CF basis is that the various CFs occurring in the full valence space are unknown. The quod erat inveniendum process by which the transformation to the unknown CFs is found yields also the CFs themselves. Each SAAP is factored to a product of atomic space-spin products, and these atomic products are compared to the atomic states given in the tables in the appendix. This comparison reveals which atomic states, or linear combinations of atomic states, occur in each SAAP.

An alternate method for the computation of T is to diagonalize the Hamiltonian matrix between SAAPs constructed from the free atoms' MBS AOs with all atoms separated by large distances ($\geq$10 Å). Each eigenvector is a pure CF, except in cases where several atomic states of the same symmetry exist, in which case this is the desired atomic near degeneracy mixing.

In the next three sections, the IACC correction is applied to H_2, NH, and F_2. For simplicity, the total antisymmetrizer A is assumed to incorporate the usual normalizing factors of $1/\sqrt{2}$ for each doubly occupied orbital occurring in a SAAP. The coset antisymmetrizer $\tilde{A}$ also includes the various $\sqrt{N!}$ factors that occur in its usual definition. Three other operators will be used in the following to simplify the

notation:

$$S = \{1 - \text{interchange of } S_z = 1/2 \text{ and } S_z = -1/2\}/\sqrt{2}$$

$$L = \{1 - \text{replacement of } L_z = +1 \text{ with } L_z = -1\}/\sqrt{2} \qquad (6.15)$$

$$X = \{1 + \text{exchange of left and right atom}\}/\sqrt{2} \quad .$$

The notation used to label each atomic state is $|^{2S+1}L(L_z,S_z)>$, which may include a superscript on the ket to indicate its charge. L_z is not a good quantum number, as described in the Appendix. The correlation correction depends only on S and L, of course, so the operators in Eq. (6.15) group only terms which receive identical corrections.

2. H_2

H_2 is the simplest possible example of the IACC correction procedure. For the ground $^1\Sigma_g^+$ state of hydrogen, the FORS wavefunction consists of the two configurations, $1\sigma_g^2$ and $1\sigma_u^2$, which correctly dissociates H_2 to 2H. Calculations at R = 1.4 bohr using an unscaled (10s,3p,1d/5s,3p,1d) basis, with molecule optimized polarization exponents of ζ_p = 0.3, 1.2, 5.4 and ζ_d = 1.96, give SCF and FORS energies of -1.1335808 h and -1.1521191 h, respectively. These correspond to bond strengths of 3.64 eV and 4.14 eV, respectively, compared to the experimental 4.75 eV (Huber and Herzberg, 1979), so that the FORS wavefunction recovers just 45% of the molecular correlation. The QBO set used here is essentially complete for SCF or FORS type calculations, with a 5 μh error in the atomic H energy, and a 15 μh error in the molecular SCF energy.

Following projected localization of the FORS orbitals, two s shaped orthonormal orbitals are obtained, one on each atom and designated A and B. One of these orbitals is shown in Fig. 6.1. It has an overlap of only 0.899 with the 1s orbital of a free H atom, due more to shrinkage of the orbital than its orthogonalizing tail on the other atom. Three singlet SAAPs can be formed from these PLOs, namely,

$$A\{AB\theta_{00}\} \quad , \quad A\{AA\theta_{00}\} \quad , \quad A\{BB\theta_{00}\} \quad , \quad \theta_{00} = (\alpha\beta - \beta\alpha)/\sqrt{2} \quad .$$

Two linear combinations of these adapted to $^1\Sigma_g^+$ symmetry can be formed,

$$|\text{covalent}\rangle = A\{AB\theta_{00}\}$$

$$|\text{ionic}\rangle = 1/\sqrt{2}\, A\{AA\theta_{00}\} + 1/\sqrt{2}\, A\{BB\theta_{00}\} \quad .$$

The ionic minus linear combination is the $^1\Sigma_u^+$ configuration $1\sigma_g^1 1\sigma_u^1$, and does not contribute to the ground state. Recomputation of the FORS function using the PLOs gives the coefficients of the covalent and ionic terms in the uncorrected FORS function as 0.780814 and 0.624763, respectively.

The covalent-ionic representation of the FORS function for H_2 is the same form as the usual valence bond function composed of nonorthogonal AOs. It differs in that the "atomic" orbitals A and B are orthonormal PLOs in the present case. As a result, the coefficient of the ionic term is considerably larger than in a conventional valence bond function. However, the FORS function is exactly equivalent to a GVB calculation in the same basis. Proof of this, and further discussion of the relationship between the MO and VB type functions for H_2, is given by Hurley (1976).

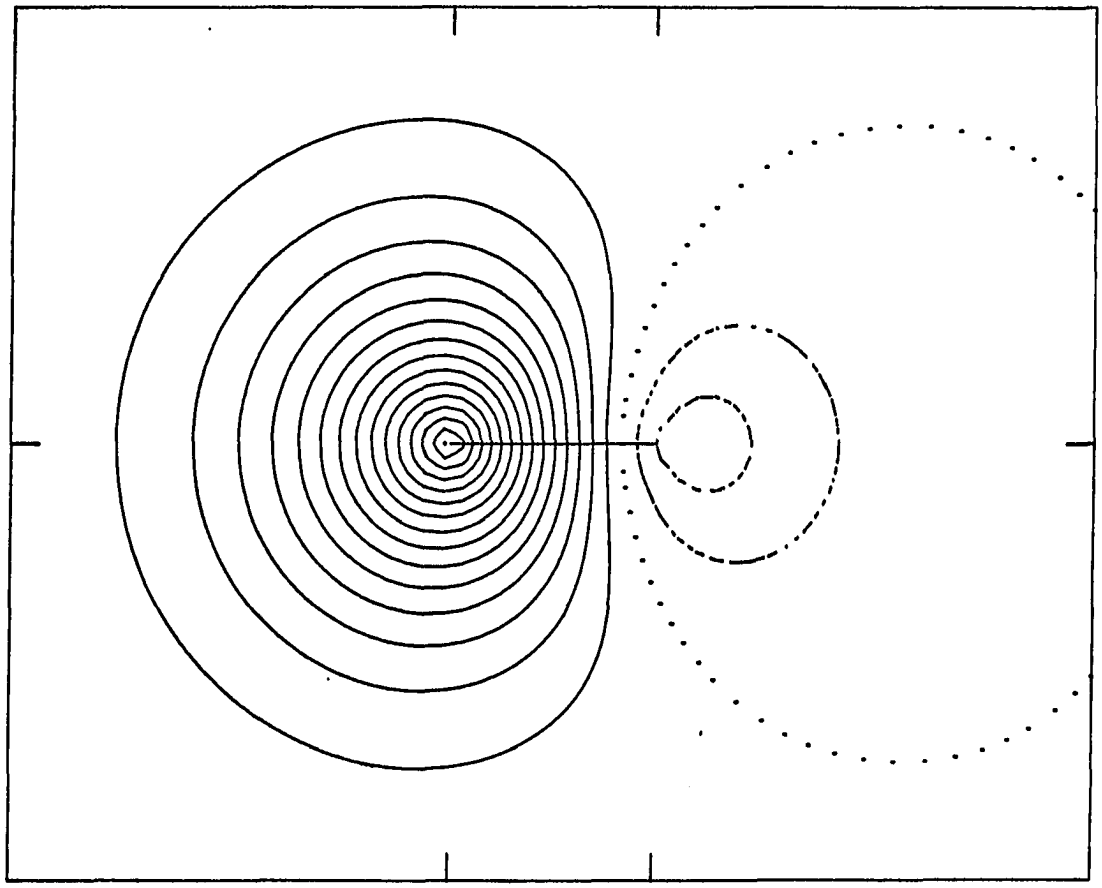


Figure 6.1. One of the equivalent projected localized orbitals for H_2. Electron occupancy = 1, bond order to other atom = 0.9756

The possible atomic states for a proton having 0, 1, or 2 electrons are

$$H^+ = |\text{proton}\rangle^+$$

$$H = |^2S(0,1/2)\rangle = 1s\alpha$$

$$H = |^2S(0,-1/2)\rangle = 1s\beta$$

$$H^- = |^1S(0,0)\rangle^- = 1s^2(\alpha\beta - \beta\alpha)/\sqrt{2} \quad .$$

One of the basis IACC postulates is to identify the molecular orbitals A and B with the atomic orbital 1s. Expanding the spin function of the covalent SAAP gives

$$|\text{covalent}\rangle = \frac{1}{\sqrt{2}} A\{(A\alpha)(B\beta)\} - \frac{1}{\sqrt{2}} A\{(A\beta)(B\alpha)\} \quad .$$

Substituting the atomic state functions into the covalent and ionic terms gives

$$|\text{covalent}\rangle = \frac{1}{\sqrt{2}} \tilde{A}\{|^2S(0,1/2)\rangle|^2S(0,-1/2)\rangle - |^2S(0,-1/2)\rangle|^2S(0,1/2)\rangle\}$$

$$|\text{ionic}\rangle = \frac{1}{\sqrt{2}} \{|^1S(0,0)\rangle^-|\text{proton}\rangle^+ + |\text{proton}\rangle^+|^1S(0,0)\rangle^-\} \quad .$$

Use of the interchange operators defined in Section VI.B.1 simplifies the notation to

$$|\text{covalent}\rangle = \tilde{A}S|^2S(0,1/2)\rangle|^2S(0,-1/2)\rangle$$

$$|\text{ionic}\rangle = X|^1S(0,0)\rangle^-|\text{proton}\rangle^+ \quad .$$

The transformation from the covalent and ionic SAAPs to the symmetry adapted CFs is thus a unit matrix. The CFs present in the molecular wavefunction are the neutral HH and the ionic H^-H^+, all in their ground states. The neutral and ionic CFs in the uncorrected FORS wavefunction have the same coefficients as the covalent and ionic SAAPs.

The IACC correction is quite simple for H_2. A bare proton and a hydrogen atom have no correlation energy. Relativistic corrections to the H and H^- energy are negligible on the chemical energy scale. Thus, only the correlation correction for H^- is required. The SCF energy of the hydride ion is -0.487927 h (Froese Fischer, 1977) and the exact energy is -0.527751 h (Pekeris, 1958). The hydride correlation correction is, therefore, -39.8 mh. The second basic assumption of the IACC scheme is that this correlation correction for H^- transfers unchanged into the hydrogen molecule. Denoting the sum of the corrections to the CFs as δ, and treating the IACC correction term in (6.12) as a perturbation, the total IACC correction is

$$\begin{aligned}
\Delta E &= C_{HH}^2\,\delta_{HH} + C_{H^-H^+}^2\,\delta_{H^-H^+} \\
&= (0.780814)^2(0 + 0) + (0.624763)^2(-0.039824 + 0) \\
&= -0.01554 \text{ h} \\
&= -0.42 \text{ eV}
\end{aligned}$$

$$E^{IACC} = E^{FORS} + \Delta E = -1.16766 \text{ h}. \quad ,$$

corresponding to D_e = 4.56 eV. This is substantially closer to the 4.75 eV experimental result than is the uncorrected FORS value of

4.14 eV. Correcting the Hamiltonian according to Eq. (6.12) and diagonalizing it gives an IACC correction of -0.43 eV, which is scarcely different from the perturbation treatment. The perturbation result is more enlightening about the nature of th IACC correction than is the diagonalization, which is not a variational treatment due to approximating the Hamiltonian. The perturbative treatment shows the IACC correction to be entirely due to the inclusion of dynamical correlation in the hydride component of the H_2 wavefunction.

3. NH

The IACC correction to the two lowest states of the imidogen radical, NH, will be illustrated. These are the $X^3\Sigma^-$ and $a^1\Delta$ states, arising from the π^2 configuration. Energy results from SCF and FORS calculations on the triplet were discussed in Chapter II, and the PLOs for this state are shown in Fig. 3.1. The nine SAAPs in the FORS wavefunction constructed from these PLOs are

$$\phi_1 = A\{s^2zhxy\ \theta_1\}$$

$$\phi_2 = A\{s^2zhxy\ \theta_3\}$$

$$\phi_3 = A\{z^2shxy\ \theta_1\}$$

$$\phi_4 = A\{z^2shxy\ \theta_3\}$$

$$\phi_5 = A\{s^2h^2xy\ \theta_1\}$$

$$\phi_6 = A\{h^2szxy\ \theta_3\}$$

$$\phi_7 = A\{h^2 szxy\ \theta_1\}$$

$$\phi_8 = A\{z^2 h^2 xy\ \theta_1\}$$

$$\phi_9 = A\{s^2 z^2 xy\ \theta_1\} \quad ,$$

where the $1s^2$ core is understood; s, x, y, and z designate the nitrogen PLOs; h is the hydrogen PLOs; and the triplet Serber spin functions are

$$\theta_1 = \theta_{00}\theta_{00}\theta_{11}$$

$$\theta_3 = \theta_{00}(\theta_{10}\theta_{11} - \theta_{11}\theta_{10})/\sqrt{2} \quad .$$

The geminal spin functions are given as footnotes to Table 6.3. The first four SAAPs correspond to NH, the next four to N^+H^-, and the last to N^-H^+.

Because the analysis of the SAAPs in terms of CFs is easiest if the two atoms have an even number of electrons, consider first one of the N^+H^- SAAPs.

$$\begin{aligned}\phi_5 &= A\{s^2h^2xy\ \theta_{00}\theta_{00}\theta_{11}\} \\ &= A\{s^2xy\ \theta_{00}\theta_{11} \cdot h^2\theta_{00}\} \\ &= \tilde{A}[A_4\{s^2xy\ \theta_{00}\theta_{11}\} \cdot A_2\{h^2\theta_{00}\}]\end{aligned}$$

Recourse to the Appendix of this chapter shows this factorization gives a ^3P state of N^+, and a singlet hydride ion,

$$\phi_5 = \tilde{A}|^3P(0,1)s^2p^2>^+|^1S(0,0)>^- \quad .$$

Here the N state is always written first in the CFs. The analysis of ϕ_8 and ϕ_9 is similar.

The ionic SAAPs ϕ_6 and ϕ_7 are more difficult to analyze, because they contain a single s electron. For example,

$$\phi_6 = A\{h^2 szxy\ \theta_{00}(\theta_{11}\theta_{10} - \theta_{10}\theta_{11})/\sqrt{2}\}$$

$$= \tilde{A}[A_4\{szxy(\theta_{11}\theta_{10} - \theta_{10}\theta_{11})/\sqrt{2}\} \cdot A_2\{h^2\ \theta_{00}\}]$$

$$= \tilde{A}[A_4\{xyzs(\theta_{11}\theta_{10} - \theta_{10}\theta_{11})/\sqrt{2}\} \cdot A_2\{h^2\ \theta_{00}\}] \quad .$$

The coupling of an s electron to a p^N configuration is treated in the Appendix to this chapter. The nitrogen ion in ϕ_6 is present in a mixture of states,

$$\phi_6 = \tilde{A}\{\frac{1}{\sqrt{3}} \left|^3 S(0,1)sp^3\right>^+ - \sqrt{\frac{2}{3}} \left|^3 D(0,1)sp^3\right>^+\} \left|^1 S(0,0)\right>^- \quad .$$

The covalent type SAAPs require careful analysis, because one of the geminal spin functions must be decomposed in separating the nitrogen electron from the hydrogen. As an example, consider

$$\phi_1 = A\{s^2 zhxy\ \theta_{00}\theta_{00}\theta_{11}\}$$

$$= A\{s^2 xyzh\ \theta_{00}\theta_{11}\theta_{00}\}$$

$$= \tilde{A}[A_5\{s^2 xyz\ \theta_{00}\theta_{11}\alpha\}\{h\beta\} - A_5\{s^2 xyz\ \theta_{00}\theta_{00}\beta\}\{h\alpha\}]/\sqrt{2} \quad .$$

The p portion of the spin function for nitrogen in the first term is the high spin Serber function $\theta^{3/2\ 3/2}$ for three unpaired electrons. However, in the second term, the spin dependence is not a single Serber

function. In fact,

$$\theta_{11}{}^{\beta} = \frac{1}{\sqrt{3}}\theta^{3/2\ 1/2} - \sqrt{\frac{2}{3}}\theta^{1/2\ 1/2} \quad ,$$

as can be shown in the same manner the Ω^{11} states are related to Serber functions in the Appendix. ϕ_1 becomes

$$\phi_1 = \tilde{A}[\frac{1}{\sqrt{2}} A_5\{s^2xyz\ \theta_{00}\theta^{3/2\ 3/2}\}\{h\beta\} + \frac{1}{\sqrt{6}} A_5\{s^2xyz\ \theta_{00}\theta^{3/2\ 1/2}\}\{h\alpha\}$$

$$- \frac{1}{\sqrt{3}} A_5\{s^2xyz\ \theta_{00}\theta^{1/2\ 1/2}\}\{h\alpha\}] \quad .$$

Use of Table 6.10 permits expression in terms of CFs,

$$\phi_1 = \tilde{A}[- \frac{1}{\sqrt{2}} \left|{}^4S(0,3/2)s^2p^3\right\rangle \left|{}^2S(0,-1/2)\right\rangle$$

$$+ \frac{1}{\sqrt{6}} \left|{}^4S(0,1/2)s^2p^3\right\rangle \left|{}^2S(0,1/2)\right\rangle - \frac{1}{\sqrt{3}} \left|{}^2S(0,1/2)s^2p^3\right\rangle \left|{}^2S(0,1/2)\right\rangle].$$

The two ${}^4S \cdot {}^2S$ products differ only in the S_z values and are treated together, with the same IACC correction.

After similar analysis of all nine SAAPs in the ground state configuration, the following nine CFs are found present:

$$\psi_1 = \tilde{A}\left|{}^4S(0,*)s^2p^3\right\rangle \left|{}^2S(0,*)\right\rangle$$

$$\psi_2 = \tilde{A}\left|{}^2D(0,1/2)s^2p^3\right\rangle \left|{}^2S(0,1/2)\right\rangle$$

$$\psi_3 = \tilde{A}\left|{}^4P(0,*)s^1p^4\right\rangle \left|{}^2S(0,*)\right\rangle$$

$$\psi_4 = \tilde{A}\left|{}^2P(0,1/2)s^1p^4\right\rangle \left|{}^2S(0,1/2)\right\rangle$$

$$\hat{\psi}_5 = \tilde{A}|^3P(0,1)s^2p^2\rangle^+|^1S(0,0)\rangle^-$$

$$\psi_6 = \tilde{A}|^3D(0,1)s^1p^3\rangle^+|^1S(0,0)\rangle^-$$

$$\psi_7 = \tilde{A}|^3S(0,1)s^1p^3\rangle^+|^1S(0,0)\rangle^-$$

$$\hat{\psi}_8 = \tilde{A}|^3P(0,1)s^0p^4\rangle^+|^1S(0,0)\rangle^-$$

$$\psi_9 = \tilde{A}|^3P(0,1)s^2p^4\rangle^-|\text{proton}\rangle^+ \quad .$$

Here the * notation accomplishes much the same purpose as the spin projection exchange operator given in Eq. (6.15) did in the H_2 example:

$$\tilde{A}|^4S(0,*)\rangle|^2S(0,*)\rangle \doteq \frac{1}{2}A|^4S(0,1/2)\rangle|^2S(0,1/2)\rangle - \frac{\sqrt{3}}{2}A|^4S(0,3/2)\rangle|^2S(0,-1/2)\rangle$$

$$\tilde{A}|^4P(0,*)\rangle|^2S(0,*)\rangle = \frac{1}{2}A|^4P(0,1/2)\rangle|^2S(0,1/2)\rangle - \frac{\sqrt{3}}{2}A|^4P(0,3/2)\rangle|^2S(0,1/2)\rangle \quad .$$

The transformation from the SAAP to CF basis is given in Table 6.1. Two of the CFs, $\hat{\psi}_5$ and $\hat{\psi}_8$, contain ^3P states, which can interact to form two configuration atomic states, from which the 2s-2p near degeneracy correlation is removed,

$$(\psi_5 \quad \psi_8) = (\hat{\psi}_5 \quad \hat{\psi}_8) \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad .$$

The coefficients a and b are fixed by a calculation on the lower ^3P atomic state of N^+, a = 0.989949 and b = -0.141422. Application of this 2 x 2 transformation following the transformation in Table 6.1 gives the

Table 6.1. Transformation from SAAPs to CFs for $X^3\Sigma^-$ NH[a]

SAAP \ CF	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6	ψ_7	ψ_8	ψ_9
ϕ_1	$\sqrt{2/3}$	$\sqrt{1/3}$							
ϕ_2	$-\sqrt{1/3}$	$\sqrt{2/3}$							
ϕ_3			$\sqrt{2/3}$	$\sqrt{1/3}$					
ϕ_4			$-\sqrt{1/3}$	$\sqrt{2/3}$					
ϕ_5					1				
ϕ_6						$-\sqrt{2/3}$	$\sqrt{1/3}$		
ϕ_7						$\sqrt{1/3}$	$\sqrt{2/3}$		
ϕ_8								1	
ϕ_9									1

[a] $\hat{\psi}_5$ and $\hat{\psi}_8$ are subjected to a further transformation to obtain the final CFs. See text for discussion.

combined transformation from the SAAP basis to the CF basis in which the IACC correction is applied.

Transforming the Hamiltonian in the SAAP basis to the CF basis as in Eq. (6.14), applying the appropriate corrections to the diagonal elements, and diagonalizing gives an IACC correction for triplet NH of -0.2476 h. First order perturbation theory gives a scarcely different correction of -0.2466 h, which is the value of the IACC correction used below. The correction at infinite separation is the $|^4S\rangle|^2S\rangle$ correction, namely -0.2148 h. The IACC correction increases the calculated bond strength to 3.65 eV, quite close to the experimental 3.85 eV.

The details of the perturbation correction for triplet NH are shown in Table 6.2. This table gives the coefficients for each CF in the uncorrected wavefunction. Although the largest term is the product of N and H in their ground states, there are substantial admixtures of other CFs, particularly the N^-H^+ CF. This CF is followed in importance by various N^*H CFs, and the N^+H^- CFs are the least important. All but one of the CFs has a larger total correction than the $|^4S\rangle|^2S\rangle$ CF, so that the correction to NH is greater than that for N + H. Because of this, the IACC bond strength is greater than the FORS bond energy.

The first excited state of NH is the $a\,^1\Delta$ state, arising from the same π^2 configuration as the ground state. The $^1\Delta$ state dissociates to the first excited state of N, ^2D, and a ground state H atom. Calculation at R = 2.0 bohr with the same basis as used for the ground state gives an SCF energy of -54.908654 h and a FORS energy of -54.930066 h. The SCF energy of ^2D N in the basis used here is -54.295505 h. The SCF

Table 6.2. IACC perturbation correction to $X^3\Sigma^-$ NH

CF	Coefficient	Nitrogen[a] Correlation	Nitrogen[a] Relativity	Hydrogen[b] Correlation	Total Correction	Weighted[c] Correction
$\|^4S\rangle\|^2S\rangle$	0.566690	-0.188[d]	-0.0268	---	-0.2148	-0.0690
$\|^2D\rangle\|^2S\rangle$	0.346512	-0.206[d]	-0.0268	---	-0.2328	-0.0280
$\|^4P\rangle\|^2S\rangle$	0.334704	-0.202	-0.0242	---	-0.2262	-0.0253
$\|^2P\rangle\|^2S\rangle$	0.161077	-0.293	-0.0242	---	-0.3172	-0.0082
$\|^3Ps^2p^2\rangle^+\|^1S\rangle^{-e}$	0.215703	-0.145[d]	-0.0270	-0.0398	-0.2118	-0.0099
$\|^3D\rangle^+\|^1S\rangle^-$	-0.164031	-0.164	-0.0238	-0.0398	-0.2276	-0.0061
$\|^3S\rangle^+\|^1S^-\rangle^-$	-0.194511	-0.200	-0.0238	-0.0398	-0.2636	-0.0100
$\|^3Ps^0p^4\rangle^+\|^1S\rangle^{-e}$	0.120096	-0.213	-0.0246	-0.0398	-0.2774	-0.0040
$\|^3P\rangle^-\|proton\rangle^+$	0.543257	-0.265[f]	-0.0266	---	-0.2916	-0.0861
						-0.2466 h

[a]Corrections from Desclaux, Moser and Verhaegen (1971), except as noted.

[b]See discussion for H_2, Section VI.B.2.

[c]Total correction for the CF, times its coefficient squared.

[d]Verhaegen and Moser (1970).

[e]The notation indicates which configuration dominates the two configuration N^+ state.

[f]From the electron affinity given by Hotop and Lineberger (1975), and data in references a and d.

and FORS bond strengths are 3.08 eV and 3.67 eV, respectively. The experimental bond strength can be obtained by appropriately combining the dissociation energy of the ground state (Piper, 1979), the electronic excitation to $^1\Delta$ and zero point energies in both states (Huber and Herzberg, 1979), and the nitrogen $^4S \rightarrow {}^2D$ excitation (Moore, 1949). This "experimental" bond strength is 4.67 eV, and possesses the ±0.10 eV uncertainty of the ground state bond energy. The FORS calculation recovers just 37% of the SCF error in the bond strength, an even lower percentage than for the ground state.

The analysis of the $a^1\Delta$ state of NH is much simpler than for the ground state. Using the $^1\Delta$ PLOs (which are much like those of the ground state), there are six SAAPs in the FORS function,

$$\phi_1 = A\{s^2zhxy\ \theta_1\}$$

$$\phi_2 = A\{z^2shxy\ \theta_1\}$$

$$\phi_3 = A\{s^2h^2xy\ \theta_1\}$$

$$\phi_4 = A\{z^2h^2xy\ \theta_1\}$$

$$\phi_5 = A\{h^2szxy\ \theta_1\}$$

$$\phi_6 = A\{s^2z^2xy\ \theta_1\} \quad ,$$

where the $1s^2$ core is understood and the singlet Serber function $\theta_1 = \theta_{00}\theta_{00}\theta_{00}$. The other possible singlet function θ_2 is not used for the last three SAAPs, as triplet coupling the π electrons gives an overall symmetry of $^1\Sigma^-$. Each of the SAAPs is equal to a single CF,

apart from a phase change resulting from sign choices in the tables in the Appendix. The CFs present are

$$\psi_1 = \tilde{A}S|^2D(-2,1/2)s^2p^3\rangle|^2S(0,-1/2)\rangle$$

$$\psi_2 = \tilde{A}S|^2D(-2,1/2)s^1p^4\rangle|^2S(0,-1/2)\rangle$$

$$\hat{\psi}_3 = \tilde{A}|^1D(-2,0)s^2p^2\rangle^+|^1S(0,0)\rangle^-$$

$$\hat{\psi}_4 = \tilde{A}|^1D(-2,0)s^0p^4\rangle^+|^1S(0,0)\rangle^-$$

$$\psi_5 = \tilde{A}|^1D(-2,0)s^1p^3\rangle^+|^1S(0\ 0)\rangle^-$$

$$\psi_6 = |^1D(-2,0)s^2p^4\rangle^-|\text{proton}\rangle^+ \quad ,$$

where the spin projection interchange operator of Eq. (6.15) has been used. There are again two CFs ψ_3 and ψ_4 which interact in a N^+ ^1D calculation. This 2s-2p near degeneracy effect is removed by an additional transformation,

$$(\psi_3 \quad \psi_4) = (\hat{\psi}_3 \quad \hat{\psi}_4) \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad ,$$

where from a two configuration calculation on the lower ^1D N^+ state, a = 0.990004 and b = -0.141041. Parity considerations perclude any mixing of ψ_1 and ψ_2, or ψ_5 with ψ_3 and ψ_4. The total transformation from SAAPs to CFs is a negative unit matrix followed by the above 2 x 2 rotation.

The first order perturbation theory IACC correction to a $^1\Delta$ NH is -0.2632 h, compared to the -0.2328 h correction for a separated hydrogen atom and ^2D nitrogen atom. The IACC bond strength is thus increased

from the FORS result to a value of 4.50 eV, not far from the true value of 4.67 eV. Explicit diagonalization of the corrected CF Hamiltonian gives an IACC correction of -0.2641 h.

The details of the IACC correction are shown in Table 6.3. The coefficients of the CFs in the uncorrected wavefunction are included in this table, and are qualitatively similar to those for the ground state. Again, the most important CF is the dissociative product closely followed by a N^-H^+ CF. N^*H CFs make the next greatest contribution, followed by N^+H^- CFs. All but one of the CFs have a greater correction than the $|^2Ds^2p^3>|^2S>$ CF, which is why singlet NH has a larger IACC correction than the dissociated hydrogen plus ^2D nitrogen.

Table 6.4 contains a summary of the relevant energetics of the N, H, and NH system. By construction, the FORS + IACC energies of the atoms are the exact energies, apart from very small basis expansion errors. The deviation of the FORS + IACC molecular energies from the exact values reflects the small (0.2 eV) errors in the bond energies for both states. Table 6.4 also shows that the vibrationless FORS + IACC separation of the two NH states is essentially the exact result. A common procedure to correct SCF or FORS excitation energies is to reduce them by the 0.49 eV error in the nitrogen $^4S - ^2D$ separation. This gives adjusted SCF and FORS excitation energies for NH of 1.34 eV and 1.49 eV. This adjusted FORS result is deceptively good, as Tables 6.2 and 6.3 clearly show the $^3\Sigma^-$ and $^1\Delta$ states are hardly pure $|^4S>|^2S>$ and $|^2D>|^2S>$.

Table 6.3. IACC perturbation correction to $a\,^1\Delta$ NH

CF	Coefficient	Nitrogen[a] Correlation	Nitrogen[a] Relativity	Hydrogen[b] Correlation	Total Correction	Weighted[c] Correction
$\|^2Ds^2p^3\rangle\|^2S\rangle$	0.663786	-0.206[d]	-0.0268	---	-0.2328	-0.1026
$\|^2Ds^1p^4\rangle\|^2S\rangle$	0.365005	-0.258	-0.0242	---	-0.2822	-0.0376
$\|^1Ds^2p^2\rangle^+\|^1S\rangle^{-e}$	0.221032	-0.157[d]	-0.0270	-0.0398	-0.2238	-0.0109
$\|^1Ds^0p^4\rangle^+\|^1S\rangle^{-e}$	0.089176	-0.248	-0.0246	-0.0398	-0.3124	-0.0025
$\|^1Ds^1p^3\rangle^+\|^1S\rangle^-$	-0.241630	-0.223	-0.0238	-0.0398	-0.2866	-0.0167
$\|^1Ds^2p^4\rangle^-\|proton\rangle^+$	0.557645	-0.272[f]	-0.0266	---	-0.2986	-0.0929
						-0.2632 h

[a]Corrections are from Desclaux, Moser and Verhaegen (1971), except as noted.

[b]See discussion for H_2, Section VI.B.2.

[c]Total correction for the CF, times its coefficient squared.

[d]Verhaegen and Moser (1970).

[e]The notation indicates which configuration dominates the two configuration N^+ state.

[f]From the electron affinity of this metastable ion given by Hotop and Lineberger (1975), and data in references a and d.

Table 6.4. Energies for N, H, and NH[a]

	SCF	FORS	FORS + IACC	Exact
N ^4S	-54.4004	-54.4004	-54.6152	-54.6158
N ^2D	-54.2955	-54.2955	-54.5283	-54.5290
T(^2D - ^4S)	2.85 eV	2.85 eV	2.36 eV	2.36 eV[b]
H ^2S	-0.4998	-0.4998	-0.4998	-0.5
NH $^3\Sigma^-$	-54.9758	-55.0027	-55.2493	-55.2573
NH $^1\Delta$	-54.9087	-54.9301	-55.1933	-55.2006
T($^1\Delta$ - $^3\Sigma^-$)	1.83 eV	1.98 eV	1.52 eV	1.54 eV[c]

[a]Energies in Hartree, except as indicated. Since correlation corrections are available only to three digits, the final digit in columns three and four is uncertain.

[b]Calculated from data in this table. The actual splitting is 2.38 eV (Moore, 1949). See note a.

[c]Calculated from data in this table. The actual splitting is 1.56 eV, correcting the observed ν_{00} for zero point vibrations (Huber and Herzberg, 1979). See note a.

4. F_2

SCF and FORS calculations on the ground state of F_2 were given in Chapter II. Plots of the PLOs for F_2 were shown in Fig. 3.3. In this section, the IACC correction is applied to F_2. Once again, the easiest way to describe the SAAPs is with a hole notation, indicating which of the PLOs are vacant. The following eight $^1\Sigma_g^+$ symmetry adapted linear combinations of SAAPs formed from PLOs are present in the FORS wavefunction:

$$\phi_1 = |z_\ell^{-1}\ z_r^{-1}\rangle$$

$$\phi_2 = \{|x_\ell^{-1}\ x_r^{-1}\rangle + |y_\ell^{-1}\ y_r^{-1}\rangle\}/\sqrt{2}$$

$$\phi_3 = \{|s_\ell^{-1}\ z_r^{-1}\rangle + |s_r^{-1}\ z_\ell^{-1}\rangle\}/\sqrt{2}$$

$$\phi_4 = |s_\ell^{-1}\ s_r^{-1}\rangle$$

$$\phi_5 = \{|z_\ell^{-2}\rangle + |z_r^{-2}\rangle\}/\sqrt{2}$$

$$\phi_6 = \{|x_\ell^{-2}\rangle + |x_r^{-2}\rangle + |y_\ell^{-2}\rangle + |y_r^{-2}\rangle\}/2$$

$$\phi_7 = \{|s_\ell^{-1}\ z_\ell^{-1}\rangle + s_r^{-1}\ z_r^{-1}\rangle\}/\sqrt{2}$$

$$\phi_8 = \{|s_\ell^{-2}\rangle + s_r^{-2}\rangle\}/\sqrt{2} \quad .$$

Here s, x, y, and z refer to the various 2s and 2p PLOs, and ℓ and r refer to the left or right atoms. The first four SAAPs are covalent, while the last four are ionic in nature. Note that the plus signs in ϕ_3 and ϕ_7 are a consequence of the phase choice for the z_ℓ and z_r

orbitals, namely that these have their plus lobes directed toward each other as shown in Fig. 3.3.

Analysis of these SAAPs in terms of atomic states gives the following CFs:

$$\psi_1 = \tilde{A}S|^2P(0,1/2)s^2p^5\rangle|^2P(0,-1/2)s^2p^5\rangle$$

$$\psi_2 = \tilde{A}LS|^2P(1,1/2)s^2p^5\rangle|^2P(1,-1/2)s^2p^5\rangle$$

$$\psi_3 = \tilde{A}XS|^2P(0,1/2)s^2p^5\rangle|^2S(0,-1/2)s^1p^6\rangle$$

$$\psi_4 = \tilde{A}S|^2S(0,1/2)s^1p^6\rangle|^2S(0,-1/2)s^1p^6\rangle$$

$$\hat{\psi}_5 = \tilde{A}X|^1S(0,0)s^2p^6\rangle^-|^1S(0,0)s^2p^4\rangle^+$$

$$\psi_6 = \tilde{A}X|^1S(0,0)s^2p^6\rangle^-|^1D(0,0)s^2p^4\rangle^+$$

$$\psi_7 = \tilde{A}X|^1S(0,0)s^2p^6\rangle^-|^1P(0,0)s^1p^5\rangle^+$$

$$\hat{\psi}_8 = \tilde{A}X|^1S(0,0)s^2p^6\rangle^-|^1S(0,0)s^0p^6\rangle^+ \quad .$$

Here the interchange operators defined in Eq. (6.15) have been used. The transformation from the SAAP basis to the CF basis is given in Table 6.5. Once again the 2s-2p near degeneracy of the ^1S F^+ states must be removed by applying the transformation

$$(\psi_5 \quad \psi_8) = (\hat{\psi}_5 \quad \hat{\psi}_8) \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

following the transformation in Table 6.5. The coefficients a and b are determined in a F^+ atomic calculation on the lower ^1S state, a = 0.981733 and b = -0.190259.

Table 6.5. Transformation from SAAPs to CFs for $X\,^1\Sigma_g^+$ F_2[a]

SAAP \ CF	ψ_1	ψ_2	ψ_3	ψ_4	$\hat{\psi}_5$	ψ_6	ψ_7	$\hat{\psi}_8$
ϕ_1	1							
ϕ_2		-1						
ϕ_3			1					
ϕ_4				1				
ϕ_5					$1/\sqrt{3}$	$-2/\sqrt{6}$		
ϕ_6					$2/\sqrt{6}$	$1/\sqrt{3}$		
ϕ_7							1	
ϕ_8								1

[a] $\hat{\psi}_5$ and $\hat{\psi}_8$ are subject to a further transformation to obtain the final CFs. See text for discussion.

The details of the IACC correction are shown in Table 6.6. The first order perturbation theory correction to F_2 is -0.8219 h, while at infinite separation, the correction to F + F in their ground states is -0.8090 h. The larger correction to the molecule than the separated atom increases the bond strength from the uncorrected FORS value. The IACC correction gives a bond strength of 1.04 eV, which is not as close to the experimental value of 1.658 eV as was the case for H_2 and NH.

The coefficients of the uncorrected FORS wavefunction in the CF basis are given in Table 6.6. The F_2 wavefunction is dominated by two neutral fluorines in their ground states. The next most important term is the ionic F^-F^+ CF, with the F^+ in its lowest singlet state, ^1D. The correction to this ionic SAAP is scarcely larger than that for the neutral ground state FF SAAP, which is responsible for the small (0.0129 h) IACC stabilization of F_2 over 2F.

Some of the corrections to the CFs given in Table 6.6 exceed 1 h. One might well wonder about the treatment of such corrections by perturbation theory, as 1 h = 627.5 kcal/mole. However, the difference in the corrections applied to any two SAAPs is smaller, at most 0.2228 h and usually much smaller. These differences in the corrections are about an order to magnitude less than the differences in the diagonal elements themselves. An even more conclusive justification of the perturbative treatment is that diagonalization of the corrected Hamiltonian gives a correction of -0.8229 h, hardly changed from the perturbative result in Table 6.6.

Table 6.6. IACC perturbation correction to $X\,^1\Sigma_g^+\ F_2$

CF	Coefficient	Left Fluorine[a] Correlation	Left Fluorine[a] Relativity	Right Fluorine[a] Correlation	Right Fluorine[a] Relativity	Total Correction	Weighted[b] Correction
$\|^2P_0\rangle\|^2P_0\rangle$	0.830507	-0.324[c]	-0.0805	-0.324[c]	-0.0805	-0.8090	-0.5580
$\|^2P_{\pm1}\rangle\|^2P_{\pm1}\rangle$	0.021155	-0.324[c]	-0.0805	-0.324[c]	-0.0805	-0.8090	-0.0004
$\|^2P_0\rangle\|^2S\rangle$	0.188533	-0.324[c]	-0.0805	-0.439	-0.0768	-0.9203	-0.0327
$\|^2S\rangle\|^2S\rangle$	0.025763	-0.439	-0.0768	-0.439	-0.0768	-1.0316	-0.0007
$\|^1S\rangle^-\|^1Ss^2p^4\rangle^{+d}$	0.236557	-0.399[e]	-0.0804	-0.248[c]	-0.0808	-0.8082	-0.0452
$\|^1S\rangle^-\|^1D\rangle^+$	-0.404721	-0.399[e]	-0.0804	-0.269[c]	-0.0808	-0.8292	-0.1358
$\|^1S\rangle^-\|^1P\rangle^+$	-0.226993	-0.399[e]	-0.0804	-0.353	-0.0765	-0.9089	-0.0468
$\|^1S\rangle^-\|^1Ss^0p^6\rangle^{+d}$	0.048149	-0.399[e]	-0.0804	-0.455[f]	-0.0736	-1.0080	-0.0023
							-0.8219 h

[a]Corrections from Desclaux, Moser and Verhaegen (1971), except as noted.

[b]Total correction for the CF, times its coefficient squared.

[c]Verhaegen and Moser (1970).

[d]The notation indicates which configuration dominates the two configuration F^+ state.

[e]From the electron affinity of F (Hotop and Lineberger, 1975) and data in reference a and c.

[f]Extrapolated from the correlation corrections in reference a for the high spin, Z = 9 $s^0p^3\ ^4S$, $s^0p^4\ ^3P$, and $s^0p^5\ ^2P$ states.

(A prodigious increase in the F_2 basis over the one used in the calculations described above improves somewhat the SCF and FORS bond strengths. Using a (18s,9p,2d/5s,4p,2d) atom centered basis, with ζ_D = 0.36 and 1.26 and a (1s,1p,1d) set of bond orbitals, with optimized exponents ζ_S = 0.87, ζ_P = 0.87, and ζ_D = 1.80 gives SCF and FORS energies of -198.770765 and -198.850875 h. These correspond to bond strengths of -1.29 and 0.89 eV, respectively. It is unlikely that further improvements in the basis can increase these values by more than 0.02 eV. Most likely the 0.11 eV improvement in the SCF and FORS results would carry over to the IACC result as well, which would leave it still rather short of the experimental value.)

5. Discussion and Conclusions

A comparison of the IACC correction for the four electronic states considered here is given in Table 6.7. In all cases, the IACC correction to the FORS wavefunction gives improved agreement with experimental results. The performance of the IACC correction is particularly encouraging for H_2 and both states of NH. However, the result for F_2 is rather poorer. This is probably because the electron correlation error in the FORS function for F_2 is more interatomic than intraatomic in nature. Each F atom brings three lone pairs to the molecule, and the interaction between these pairs is not accounted for by the FORS function (which principally correlates the bond pair) or the IACC correction.

It is difficult to compare previous Atoms in Molecules work to the present IACC scheme. Most of the earlier work was done with rather

Table 6.7. Comparison of bond energies[a]

	H_2 $X^1\Sigma_g^+$	NH $X^3\Sigma^-$	NH $a^1\Delta$	F_2 $X^1\Sigma_g^+$
SCF	3.64	1.06	3.08	-1.40
FORS	4.14	2.77	3.67	0.78
FORS + IACC	4.56	3.65	4.50	1.04
exp	4.75[b]	3.85[c]	4.67[b,c,d]	1.658[b]

[a]All energies in eV.

[b]Huber and Herzberg (1979).

[c]Piper (1979).

[d]Moore (1949).

small, nonorthogonal atomic bases, and usually with a subset of the full valence space CF sets used here. No comparison to earlier work will be given, but references to earlier H_2, NH, and F_2 calculations will be given. References to work on other molecules can be found in the review articles listed above. Hurley (1955) applied the AIM method of Moffitt (1951b) to H_2, and based on its great overestimate of the bond strength proposed the ICC correction scheme (1956). NH has been frequently considered by various workers. Hurley (1958a, 1958b) has given ICC corrections to the ground and excited states of all second period hydrides, including NH. Krauss and Wehner (1958) also apply the ICC correction to NH. This paper deals with the problem of transforming a basis of Slater determinants constructed from orthonormal MOs into a valence bond form, a problem solved here with the aid of the PLOs. Liu and Verhaegen (1970) have utilized Mulliken population analyses to identify atomic configurations in the NH wavefunction. Finally, Balint-Kurti and Karplus (1969) have applied their Orthogonalized Moffitt modification of AIM theory to F_2.

A preliminary report of the IACC correction scheme for F_2 has been given by Ruedenberg et al. (1982). In this sketch of the IACC method, relativistic corrections were omitted, and basis errors for the various atomic states were included in the corrections. While relativity contributes little to the F_2 bond strength (see below), the inclusion of basis errors, particularly for the F^- ion, does affect the bond strength, giving a value of 1.30 eV. These basis corrections to the atomic states have not been applied here for the following reason. Adding

diffuse functions to the diatom basis set will contribute very little to the uncorrected FORS results. These diffuse functions will substantially reduce the basis errors in SCF calculations on the negative ions to about the same size as those for the neutral or positive ions. If basis errors are excluded from the IACC corrections, adding the diffuse functions will have almost no effect on the magnitude of the IACC correction, rather than substantially reducing the total IACC correction. Assuming atomic correlation corrections to be independent of the atomic basis expansion chosen is in keeping with the assumption by Hurley that the ICC correction is independent of orbital scaling, and the assumption in the present case that the correlation correction doesn't change when the AOs of a particular state are replaced by the molecular PLOs. There are a few sign changes in the present case from the earlier presentation, due to different phase choices in the tables in the appendix to this chapter.

An interesting byproduct of the IACC scheme is its ability to incorporate relativistic as well as correlation errors in the correction terms. Applying the perturbative treatment to the relativistic corrections only gives a total relativistic correction of -0.0262 h to $^3\Sigma^-$ NH, and -0.0262 h to $^1\Delta$ NH, compared to -0.0268 h for the dissociative limit of both states. Relativity thus decreases these bond strengths by less than 0.02 eV, and leaves the excitation energy unchanged. For F_2, the relativistic correction is -0.1607, compared to -0.1610 for 2F, which decreases the bond strength by less than 0.01 eV. Although the relativistic corrections are substantial, in

the case of F_2 and NH they are very nearly canceling. Such would not be the case if the species being compared possess greatly different occupation of the 2s orbital, as Feller and Davidson (1980) suggest may be the case for singlet and triplet methylene.

In conclusion, the full valence space FORS wavefunction may be expressed in a basis of composite functions, that is, products of neutral and ionic atomic states. The expansion of the FORS wavefunction in these CF bases reveals much about the participation of atoms and ions in the molecular wavefunction. In addition, such expansions permit the application of the IntraAtomic Correlation Correction to the molecular FORS wavefunction. This correction utilizes experimental data for the various atomic states to correct the molecular wavefunctions for dynamical correlation effects not incorporated in the FORS model. The IACC scheme gives very good results for H_2 and two states of NH, and somewhat improved results for F_2. Relativistic contributions to bond strengths are found to be negligible for the electronic states considered here.

C. Appendix: Real Atomic States

A prerequisite for IACC calculations or molecules containing the elements from the second period is the knowledge of these atom's valence states. These valence states must be expressed in terms of real atomic orbitals, as molecular calculations are performed with real PLOs. An L-S coupling scheme is used for the atomic states, except that the restriction to real orbitals precludes the use of L_z as a good quantum number.

Salmon and Ruedenberg (1972) have described a computer algorithm for the generation of atomic states that are eigenfunctions of L^2, L_z, S^2, and S_z. Karwowski and Fraga (1974), using a modification of this procedure, have tabulated all atomic states arising from the p^N, N = 2,3; d^N, N = 2,3,4,5; and f^N, N = 2,3,4,5,6,7 configurations. Both of these procedures use complex atomic orbitals, with Serber spin dependence. These algorithms rely on simultaneous diagonalization of the L^2, L_z, S^2, and S_z matrices, so that the eigenvectors, that is, the atomic states, have arbitrary phases.

1. p^N states

In the present case, the p^N atomic states are obtained using angular momentum ladder operators (Gray and Wills, 1931). In this procedure, the state with highest total angular momentum, orbital angular momentum projection, and spin angular momentum projection is written in terms of complex orbitals. The lowering operators L^- and S^- are applied to generate the states with lower projections. States with lower total orbital angular momentum are generated by orthogonality to the higher angular momentum states. The phases of the L-S states possessing the maximal projections $L_z = L$ and $S_z = S$ have arbitrary phase, but the phases of all other states are fixed by the lowering operators.

Once the atomic states constructed from complex p orbitals are found, the atomic states using real p orbitals can be formed. Within the Condon-Shortley phase convention, the complex and real p orbitals are related by the unitary transformation,

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & -i/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -i/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

Atomic states constructed from real p orbitals are found by substituting for the complex orbitals. The states with $L_z = 0$ are always pure real or pure imaginary functions of the real p orbitals. If they are imaginary, real valued atomic states are generated by multiplying by i. The pair of states formed from real atomic orbitals with orbital angular momentum $+L_z$ and $-L_z$, with all other quantum numbers equal, are either complex conjugates or negative complex conjugates. The normalized sums and differences of these pairs are thus pure real or pure imaginary functions. The pure imaginary functions are again multiplied by i to generate real states.

All real valued p^N states constructed from real p orbitals have been found by this process. The symbol L_z is used below to designate the real linear combinations of $+L_z$ and $-L_z$ eigenstates, while $-L_z$ designates the linear combination made real by multiplication by i. L_z is not a good quantum number, except when $L_z = 0$. The real atomic p^N states are given in Tables 6.8 to 6.11. In these tables, each state is labeled as $|^{2S+1}L(L_z,S_z)>$. The geminal spin functions are labeled θ_{SS_z}, and are given as footnotes to Table 6.9. The multielectron Serber spin functions are labeled $\theta_{\alpha}^{SS_z}$, where α indexes the function in cases where more than one such function is possible.

Table 6.8. p^1, p^5, p^6 states

p^1 states			
Doublets[a]	$x\alpha$	$y\alpha$	$z\alpha$
$\|^2P(1,1/2)\rangle$	-1		
$\|^2P(0,1/2)\rangle$			1
$\|^2P(-1,1/2)\rangle$		1	

p^5 states			
Doublets[b]	$y^2z^2x\ \theta_{00}\theta_{00}\alpha$	$x^2z^2y\ \theta_{00}\theta_{00}\alpha$	$x^2y^2z\ \theta_{00}\theta_{00}\alpha$
$\|^2P(1,1/2)\rangle$	-1		
$\|^2P(0,1/2)\rangle$			1
$\|^2P(-1,1/2)\rangle$		1	

p^6 state
$\|^1S(0,0)\rangle = x^2y^2z^2\ \theta_{00}\theta_{00}\theta_{00}$, $\theta_{00} = (\alpha\beta-\beta\alpha)/\sqrt{2}$

[a]Doublets with $s_z = -1/2$ are obtained by replacing α with β.

[b]Doublets with $s_z = -1/2$ are obtained by replacing $\theta_{00}\theta_{00}\alpha$ by $\theta_{00}\theta_{00}\beta$. $\theta_{00} = (\alpha\beta-\beta\alpha)/\sqrt{2}$.

Table 6.9. p^2 states

Singlets[a]	$x^2\,\theta_{00}$	$y^2\,\theta_{00}$	$z^2\,\theta_{00}$	$xy\,\theta_{00}$	$xz\,\theta_{00}$	$yz\,\theta_{00}$
$\|^1S(0,0)\rangle$	$1/\sqrt{3}$	$1/\sqrt{3}$	$1/\sqrt{3}$			
$\|^1D(2,0)\rangle$	$1/\sqrt{2}$	$-1/\sqrt{2}$				
$\|^1D(1,0)\rangle$					-1	
$\|^1D(0,0)\rangle$	$1/\sqrt{6}$	$1/\sqrt{6}$	$-2/\sqrt{6}$			
$\|^1D(-1,0)\rangle$						1
$\|^1D(-2,0)\rangle$				-1		

Triplets[b]	$xy\,\theta_{11}$	$xz\,\theta_{11}$	$yz\,\theta_{11}$
$^3P(1,1)$		-1	
$^3P(0,1)$	1		
$^3P(-1,1)$			1

[a] $\theta_{00} = (\alpha\beta-\beta\alpha)/\sqrt{2}$.

[b] Triplets with $s_z = 0$ or -1 are obtained by replacing θ_{11} by θ_{10} or θ_{1-1}, respectively. $\theta_{11} = \alpha\alpha$, $\theta_{10} = (\alpha\beta+\beta\alpha)/\sqrt{2}$, and $\theta_{1-1} = \beta\beta$.

Table 6.10. p^3 states

Doublets[a]	$x^2y\ \theta_1^{1/2\ 1/2}$	$x^2z\ \theta_1^{1/2\ 1/2}$	$y^2x\ \theta_1^{1/2\ 1/2}$	$y^2z\ \theta_1^{1/2\ 1/2}$
$\|^2P(1,1/2)\rangle$			$-1/\sqrt{2}$	
$\|^2P(0,1/2)\rangle$		$1/\sqrt{2}$		$1/\sqrt{2}$
$\|^2P(-1,1/2)\rangle$	$1/\sqrt{2}$			
$\|^2D(2,1/2)\rangle$		$1/\sqrt{2}$		$-1/\sqrt{2}$
$\|^2D(1,1/2)\rangle$			$-1/\sqrt{2}$	
$\|^2D(0,1/2)\rangle$				
$\|^2D(-1,1/2)\rangle$	$-1/\sqrt{2}$			
$\|^2D(-2,1/2)\rangle$				

Quartets[b]	$xyz\ \theta^{3/2\ 3/2}$
$\|^4S(0,3/2)\rangle$	-1

[a]The doublet spin functions are $\theta_1^{1/2\ 1/2} = \theta_{00}\alpha$, and $\theta_2^{1/2\ 1/2} = (\theta_{10}\alpha - \sqrt{2}\ \theta_{11}\beta)/\sqrt{3}$. Doublet states with $s_z = -1/2$ are obtained by replacing $\theta_1^{1/2\ 1/2}$ with $\theta_1^{1/2-1/2} = \theta_{00}\beta$ and $\theta_2^{1/2\ 1/2}$ with $\theta_2^{1/2-1/2} = (\theta_{10}\beta - \sqrt{2}\ \theta_{1-1}\alpha)/\sqrt{3}$.

[b]The quartet state with projection $s_z = 1/2$, $-1/2$, or $-3/2$ is obtained by replacing $\theta^{3/2\ 3/2} = \theta_{11}\alpha$ with $\theta^{3/2\ 1/2} = (\theta_{11}\beta + \sqrt{2}\ \theta_{10}\alpha)/\sqrt{3}$ or $\theta^{3/2-1/2} = (\theta_{1-1}\alpha + \sqrt{2}\ \theta_{10}\beta)/\sqrt{3}$ or $\theta^{3/2-3/2} = \theta_{1-1}\beta$, respectively.

$z^2x\ \theta_1^{1/2\ 1/2}$	$z^2y\ \theta_1^{1/2\ 1/2}$	$xyz\ \theta_1^{1/2\ 1/2}$	$xyz\ \theta_2^{1/2\ 1/2}$
$-1/\sqrt{2}$			
	$1/\sqrt{2}$		
$1/\sqrt{2}$			
			1
	$1/\sqrt{2}$		
		-1	

Table 6.11. p^4 states

Singlets	$x^2y^2\,\theta_{00}\theta_{00}$	$x^2z^2\,\theta_{00}\theta_{00}$	$y^2z^2\,\theta_{00}\theta_{00}$	$x^2yz\,\theta_{00}\theta_{00}$	$y^2xz\,\theta_{00}\theta_{00}$	$z^2xy\,\theta_{00}\theta_{00}$
$\|^1S(0,0)\rangle$	$1/\sqrt{3}$	$1/\sqrt{3}$	$1/\sqrt{3}$			
$\|^1D(2,0)\rangle$		$1/\sqrt{2}$	$-1/\sqrt{2}$			
$\|^1D(1,0)\rangle$					-1	
$\|^1D(0,0)\rangle$	$-2/\sqrt{6}$	$1/\sqrt{6}$	$1/\sqrt{6}$			
$\|^1D(-1,0)\rangle$				1		
$\|^1D(-2,0)\rangle$						-1

Triplets[a]	$x^2yz\,\theta_{00}\theta_{11}$	$y^2xz\,\theta_{00}\theta_{11}$	$z^2xy\,\theta_{00}\theta_{11}$
$\|^3P(1,1)\rangle$		-1	
$\|^3P(0,1)\rangle$			1
$\|^3P(-1,1)\rangle$	1		

[a]Triplets with $s_z = 0$ or -1 are obtained by replacing $\theta_{00}\theta_{11}$ with $\theta_{00}\theta_{10}$ or $\theta_{00}\theta_{1-1}$, respectively.

2. s^2p^N and s^1p^N states

The s^2p^N states are trivially generated from the p^N states given above. The space part of the appropriate p^N state is prefixed with s^2, and its Serber function is prefixed with the geminal function θ_{00}.

The s^1p^N states are most readily generated from the p^N states by coupling the s electron at the end, i.e., p^Ns^1. It is readily shown that the L and L_z quantum numbers of a p^Ns^1 state are exactly the same as the corresponding p^N state. The p^Ns^1 state is obtained by coupling an $s\alpha$ and/or an $s\beta$ electron to the appropriate p^N states.

Formulae for coupling one electron to a Serber function are given by Salmon (1974). The process is illustrated for one case with N = 3 that occurs in the NH wavefunction. Three possible states for p^3 with space product xyz and $L_z = 0$ are the $|^2D(0,1/2)>$, $|^4S(0,3/2)>$, and $|^4S(0,1/2)>$ states. The possible triplet states for p^3s^1 with space product xyzs and $L_z = 0$ and $S_z = 1$ are $|^3D(0,1)>$ and $|^3S(0,1)>$. The $|^3D(0,1)>$ state can be obtained by coupling $s\alpha$ to $|^2D(0,1/2)>$, while the $|^3S(0,1)>$ state can be reached by coupling $s\alpha$ to $|^4S(0,1/2)>$ or $s\beta$ to $|^4S(0,3/2)>$. Using Salmon's formulae, the spin dependence for $|^3D(0,1)>$ is

$$\Omega_1^{11}(4e^-) = \theta_2^{1/2\ 1/2}(3e^-)\alpha \quad .$$

For the $|^3S(0,1)>$ state, the spin function is

$$\Omega_2^{11}(4e^-) = -\frac{1}{2}\theta^{3/2\ 1/2}(3e^-)\alpha + \frac{\sqrt{3}}{2}\theta^{3/2\ 3/2}(3e^-)\beta \quad .$$

The desired p^3s^1 states are, therefore,

$$|^3D(0,1)> = A\{xyzs\ \Omega_1^{11}\}$$

$$|^3S(0,1)> = -A\{xyzs\ \Omega_2^{11}\} \quad .$$

(There is a third possible triplet function Ω_3^{11}, obtained by α coupling an electron to $\theta_1^{1/2\ 1/2}(3e^-)$, that is used in the $|^3D(-2,1)>$ state.)

In general, the spin functions $\Omega_\alpha^{SS_z}$, obtained by coupling a single electron to the Serber functions for one fewer electron, are not themselves Serber functions, but are related to them by an orthogonal transformation,

$$\Omega^{SS_z} = \theta^{SS_z} D \quad , \quad D_{\gamma\kappa} = <\theta_\gamma^{SS_z}|\theta_\kappa^{SS_z}> \quad .$$

Elements of the transformation D are most readily calculated by expanding both θ and Ω in terms of explicit α/β products. In the present case, the Serber triplet functions with $S_z = 1$ for four electrons are

$$\theta_1^{11} = \theta_{00}\theta_{11}$$

$$\theta_2^{11} = \theta_{11}\theta_{00}$$

$$\theta_3^{11} = (\theta_{11}\theta_{10} - \theta_{10}\theta_{11})/\sqrt{2} \quad .$$

Expressing the Ω spin functions in terms of these gives

$$\Omega_1^{11} = (\theta_2^{11} - \sqrt{2}\,\theta_3^{11})/\sqrt{3}$$

$$\Omega_2^{11} = (\sqrt{2}\,\theta_2^{11} + \theta_3^{11})/\sqrt{3}$$

$$\Omega_3^{11} = \theta_1^{11} \quad .$$

The p^3s^1 ^3S and ^3D states with spin projection 1 can now be expressed in terms of SAAPs, with Serber spin dependence rather than Ω's, as

triplets	xyzs θ_1^{11}	xyzs θ_2^{11}	xyzs θ_3^{11}
$\|^3D(-2,1)\rangle$	1		
$\|^3D(0,1)\rangle$		$1/\sqrt{3}$	$-\sqrt{2/3}$
$\|^3S(0,1)\rangle$		$\sqrt{2/3}$	$1/\sqrt{3}$

VII. REFERENCES

T. Arai, Rev. Mod. Phys. 32, 370 (1960).

G. G. Balint-Kurti and M. Karplus, in Orbital Theories of Molecules and Solids, edited by N. H. Marsh (Clarendon Press, Oxford, 1974), p. 250.

G. G. Balint-Kurti and M. Karplus, J. Chem. Phys. 50, 478 (1969).

R. D. Bardo and K. Ruedenberg, J. Chem. Phys. 60, 918 (1974).

N. L. Bauld, J. A. Thompson, C. E. Hudson and P. S. Bailey, J. Am. Chem. Soc. 90, 1822 (1968).

C. W. Bauschlicher, J. Chem. Phys. 72, 880 (1980).

A. J. Bellamy, J. Chem. Soc. Perkin I, 342 (1972).

F. P. Billingsley and M. Krauss, J. Chem. Phys. 60, 2767 (1974).

J. S. Binkley, J. A. Pople and W. J. Hehre, J. Am. Chem. Soc. 102, 939 (1979).

F. W. Bobrowicz and W. A. Goddard, in Methods of Electronic Structure Theory, edited by H. F. Schaefer (Plenum Press, New York, 1977), p. 79 and references therein.

S. F. Boys, Proc. R. Soc. London Ser. A 200, 542 (1950).

P. Cade and W. Huo, At. Nuc. Data Tables 15, 1 (1975).

P. Cade and A. C. Wahl, At. Nuc. Data Tables 13, 339 (1974).

J. Catalan, F. Escudero, J. Laso, O. Mo and M. Yanez, J. Mol. Struc. 69, 217 (1980).

L. M. Cheung, K. R. Sundberg and K. Ruedenberg, Int. J. of Quantum Chem. 16, 1103 (1979).

R. Cimiraglia, T.-K. Ha and Hs. H. Gunthard, Chem. Phys. Lett. 85, 262 (1982).

E. Clementi, IBM J. Res. Dev. 9, 2 (1965).

E. Clementi and H. Popkie, J. Chem. Phys. 57, 4870 (1972).

E. Clementi and C. Roetti, Atomic Data and Nuclear Data Tables 14, 177 (1974).

R. Criegee, Record Chem. Prog. 18, 111 (1957).

G. Das and A. C. Wahl, J. Chem. Phys. 44, 87 (1966).

K. L. Demerjian, J. A. Kerr and J. G. Calvert, Adv. Environ. Sci. Technol. 4, 1 (1974).

J. P. Desclaux, C. M. Moser and G. Verhaegen, J. Phys. B 4, 296 (1971).

W. von E. Doering and J. W. Rosenthal, J. Am. Chem. Soc. 89, 4534 (1967).

M. G. Dombek, Ph.D. thesis, Iowa State University, 1977 (unpublished).

T. H. Dunning, D. C. Cartwright, W. J. Hunt, P. J. Hay and F. W. Bobrowicz, J. Chem. Phys. 64, 4755 (1976).

T. H. Dunning and P. J. Hay, in Methods of Electronic Structure Theory, edited by H. F. Schaefer (Plenum Press, New York, 1977), p. 1 and references therein.

M. Dupuis, J. Chem. Phys. 74, 5758 (1981).

M. Dupuis, D. Spangler and J. J. Wendoloski, in NRCC Software Catalog (University of California, Berkeley, 1980), Vol. 1.

C. Edmiston and K. Ruedenberg, Rev. Mod. Phys 35, 457 (1963).

C. Edmiston and K. Ruedenberg, J. Chem. Phys. 43, S97 (1965).

C. Edmiston and K. Ruedenberg, Topics in Current Chemistry 23, 31 (1971).

S. T. Elbert, in Numerical Algorithms in Chemistry: Algebraic Methods, edited by C. Moler and I. Shavitt (University of California, Berkeley, 1978), p. 129.

D. F. Feller, Ph.D. thesis, Iowa State University, 1979 (unpublished).

D. Feller and E. R. Davidson, Chem. Phys. Lett. 69, 201 (1980).

D. F. Feller and K. Ruedenberg, Theoret. Chim. Acta (Berl.) 52, 231 (1979).

D. F. Feller, M. W. Schmidt and K. Ruedenberg, J. Am. Chem. Soc. 104, 960 (1982).

M. C. Flanigan, A. Kormonicki and J. W. McIver, in Semiempirical Methods of Electronic Structure Calculations, Part B: Applications, edited by G. A. Segel (Plenum Press, New York, 1977), p. 1 and references therein.

J. M. Foster and S. F. Boys, Rev. Mod. Phys. 32, 300 (1960).

S. Fraga, J. Karwowski and K. M. S. Saxena, Handbook of Atomic Data (Elsevier Scientific Publishing Co., Amsterdam, 1976).

C. Froese Fischer, The Hartree-Fock Method for Atoms (Wiley and Sons, New York, 1977), p. 165.

W. A. Goddard, J. Am. Chem. Soc. 94, 793 (1972).

N. M. Gray and L. A. Wills, Phys. Rev. 38, 248 (1931).

S. Green, J. Chem. Phys. 52, 3100 (1970).

S. Green, J. Chem. Phys. 57, 4694 (1972a).

S. Green, Chem Phys. Lett. 13, 552 (1972b).

T.-K. Ha, H. Kuhne, S. Vacanni and Hs. H. Gunthard, Chem. Phys. Lett. 24, 172 (1974).

J.-P. Hagenbuch, B. Stampfli and P. Vogel, J. Am. Chem. Soc. 103, 3934 (1981).

L. B. Harding and W. A. Goddard, J. Am. Chem. Soc. 100, 7180 (1978).

C. Harries, Chem. Ber. 36, 1933 (1903).

P. J. Hay and T. H. Dunning, J. Chem. Phys. 67, 2290 (1977).

G. Herzberg, Electronic Spectra of Polyatomic Molecules (Van Nostrand Reinhold, New York, 1966).

P. C. Hiberty, J. Am. Chem. Soc. 98, 6088 (1976).

H. Hotop and W. C. Lineberger, J. Phys. Chem. Ref. Data 4, 539 (1975).

K. P. Huber and G. Herzberg, Constants of Diatomic Molecules (Van Nostrand Reinhold, New York, 1979).

S. Hunig, H. R. Muller and W. Their, Angew. Chem., Int. Ed. Engl. 4, 271 (1965).

W. J. Hunt, P. J. Hay and W. A. Goddard, J Chem. Phys. 57, 738 (1972).

W. M. Huo, J. Chem. Phys. 43, 624 (1965).

A. C. Hurley, Proc. Phys. Soc. A 68, 149 (1955).

A. C. Hurley, Proc. Phys. Soc. A 69, 49 (1956).

A. C. Hurley, Proc. Roy. Soc. A 248, 119 (1958a).

A. C. Hurley, Proc. Roy. Soc. A 249, 402 (1958b).

A. C. Hurley, Rev. Mod. Phys. 35, 448 (1963).

A. C. Hurley, Introduction to the Electron Theory of Small Molecules (Academic Press, London, 1976).

R. P. Johnson and M. W. Schmidt, J. Am. Chem. Soc. 103, 3244 (1981).

G. Karlstrom, S. Engstrom and B. Jonsson, Chem. Phys. Lett. 67, 343 (1979).

J. Karwowski and S. Fraga, State Functions for Many Electron Atoms - Eigenfunctions of L^2 and S^2, Technical Report TC-SF-I-74 (University of Alberta, 1974).

K. Kirby-Docken and B. Liu, J. Chem. Phys. 66, 4309 (1977).

M. Krauss and J. F. Wehner, J. Chem. Phys. 29, 1287 (1958).

K. Kuchitsu, J. Chem. Phys. 44, 906 (1966).

H. Kuhne, S. Vaccani, T.-K. Ha and Hs. H. Gunthard, Chem. Phys. Lett. 38, 449 (1976).

B. Lam, SAAP program, Iowa State University, 1982 (unpublished).

G. C. Lie and E. Clementi, J. Chem. Phys. 60, 1275 and 1288 (1974).

J. Lieven, J. Breulet and G. Verhaegen, Theoret. Chim. Acta (Berlin) 60, 339 (1981).

H. P. D. Liu and G. Verhaegen, J. Chem. Phys. 53, 735 (1970).

F. J. Lovas and R. D. Suenram, Chem. Phys. Lett. 51, 453 (1977).

P. O. Löwdin, J. Chem. Phys. 18, 365 (1950).

P. O. Löwdin, Phys. Rev. 97, 1474, 1490 and 1509 (1955).

K. Mackenzie, J. Chem. Soc., 4646 (1965).

K. Mackenzie, J. Chem. Soc. (C), 1784 (1969).

R. I. Martinez, R. E. Huie and J. T. Herron, Chem. Phys. Lett. 51, 457 (1977).

C. E. Miller, J. Chem. Educ. 42, 254 (1965).

K. J. Miller and K. Ruedenberg, J. Chem. Phys. 48, 3414 (1968).

W. Moffitt, Proc. Roy. Soc. A 210, 224 (1951a).

W. Moffitt, Proc. Roy. Soc. A 210, 245 (1951b).

W. Moffitt, Rept. Prog. Phys. 17, 173 (1954).

C. E. Moore, National Bureau of Standards Monograph, 467 (1949).

R. S. Mulliken, J. Chem. Phys. 23, 1833 and 1841 (1955).

D. J. Pasto and D. M. Chipman, J. Am. Chem. Soc. 101, 2290 (1979).

C. L. Pekeris, Phys. Rev. 112, 1649 (1958).

L. G. Piper, J. Chem. Phys. 70, 3417 (1979).

R. C. Raffenetti, J. Chem. Phys. 58, 4452 (1973).

R. C. Raffenetti, Int. J. Quantum Chem. Symp. 9, 289 (1975).

J. W. Raymonda, J. S. Muenter and W. A. Klemperer, J. Chem. Phys. 52, 3458 (1970).

B. O. Roos and P. E. M. Sieghban, J. Am. Chem. Soc. 99, 7716 (1977).

B. O. Roos, P. R. Taylor and P. E. M. Siegbahn, Chem. Phys. 48, 157 (1980).

K. Ruedenberg, Phys. Rev. Lett. 27, 1105 (1971).

K. Ruedenberg, L. M. Cheung and S. T. Elbert, Int. J. Quantum Chem. 16, 1069 (1979).

K. Ruedenberg, M. W. Schmidt and M. Gilbert, Chem. Phys., in press (1982)

K. Ruedenberg and R. Poshusta, Adv. Quantum Chem. 6, 267 (1972).

K. Ruedenberg, R. C. Raffenetti and R. D Bardo, in Energy, Structure, and Reactivity, Proceedings of the 1972 Boulder Conference on Theoretical Chemistry, edited by D. W. Smith and W. B. McRae (Wiley, New York, 1973), p. 164.

K. Ruedenberg and K. R. Sundberg, in Quantum Science, edited by J. L. Calais, O. Goscinski, J. Linderberg and Y. Ohrn (Plenum Press, New York, 1976), p. 505.

R. R. Rye and R. S. Hansen, J. Phys. Chem. 73, 1667 (1969).

W. I. Salmon, in Advances in Quantum Chemistry, Vol. 8, edited by P. Lowdin (Academic Press, New York, 1974), p. 37.

W. I. Salmon, L. M. Cheung and K. Ruedenberg, J. Chem. Phys. 57, 2776 and 2787 (1972).

W. I. Salmon and K. Ruedenberg, J. Chem. Phys. 57, 2791 (1972).

H. F. Schaefer and F. E. Harris, J. Chem. Phys. 48, 4946 (1968).

H. F. Schaefer, R. A. Klemm and F. E. Harris, J. Chem. Phys. 51, 4643 (1969).

M. W. Schmidt and K. Ruedenberg, J. Chem. Phys. 71, 3951 (1979).

P. E. M. Sieghbahn, J. Chem. Phys. 70, 5391 (1979).

P. Siegbahn, A. Heiberg, B. Roos and B. Levy, Phys. Scripta 21, 323 (1980).

D. M. Silver, E. L. Mehler and K. Ruedenberg, J. Chem. Phys. 52, 1174, 1181 and 1206 (1970).

R. D. Suenram and F. J. Lovas, J. Am. Chem. Soc. 100, 5117 (1978).

K. R. Sundberg, Ph.D. thesis, Iowa State University, 1975 (unpublished).

O. M. Uy and J. Drowart, High Temp. Sci. 2, 293 (1970).

P. Valtazanos, DIAPOT program, Iowa State University, 1980 (unpublished).

G. Verhaegen and C. M. Moser, J. Phys. B 3, 478 (1970).

S. K. Vidyarthi, C. Willis, R. A. Back and R. M. McKitrick, J. Am. Chem. Soc. 96, 7647 (1974).

W. R. Wadt and W. A. Goddard, J. Am. Chem. Soc. 97, 3004 (1975).

D. D. Wagman, National Bureau of Standards Technical Note 270-3, 1968.

A. C. Wahl, J. Chem. Phys. 41, 2600 (1964).

A. C. Wahl and G. Das, in Methods of Electronic Structure Theory, edited by H. F. Schaefer (Plenum Press, New York, 1977), p. 51 and references therein.

C. Willis, R. A. Back, J. M. Parsons and J. G. Purdon, J. Am. Chem. Soc. 99, 4451 (1977).

R. B. Woodward and R. Hoffmann, The Conservation of Orbital Symmetry (Verlag Chemie, Weinheim, 1971), p. 141.

K. Yamaguchi, S. Yabushita, T. Fueno, S. Kato, K. Morokuma and S. Iwata, Chem. Phys. Lett. 71, 563 (1980).

B. Zurawski and W. Kutzelnigg, J. Am. Chem. Soc. 100, 2654 (1978).

VIII. ACKNOWLEDGEMENTS

A number of people have contributed greatly to the preparation of this dissertation. First and foremost is Professor Klaus Ruedenberg, who guided this research into hopefully useful directions. A large number of computer programs were required for this work, and the programming skills of various former members of the quantum chemistry group, particularly Richard Raffenetti, whom I have never met, and Lap Cheung provided the major programs with which this work was done. Steven Elbert provided many improvements to these programs, greatly increasing their chemical capabilities, and kept bugs to a minimum. David Feller impressed on me the need to understand the chemical significance and accuracy of computational methods. Brenda Lam proofread much of the manuscript, but made coffee all too infrequently. Lesley Swope typed the manuscript swiftly and accurately. Finally, thank you to Mark Gordon, who waited patiently.